


```
Db 432 LysGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaAspGlyIleAsp 451
QY 1212 TCCAAAGTTGATGGTATCATCATCTGACCGCCGACAGGCGCCAAATTCAGGCTCAGGGA 1271
Db 452 AlalysGlnArgMetValIleIleSerGlyProProGluAlaGlnPheLysAlaGlnGly 471
QY 1272 AGAATCTATGCAACCTCAGGAGGAGAACTTCTTGGTCCCAAGGAGGAAGTCAAGCTG 1331
Db 472 ArgIlePheGlyLysLeuLysGlnGluAsnPheGlyProLysGluGluValLysLeu 491
QY 1332 GAGACCCACATACGTGTGCCAGCAGTCAGCAGCTGGCGGGTCAATTTGGCAAGAGTGGAAAA 1391
Db 492 GluAlaHisIleLysValProSerPheAlaGlyArgValIleGlyLysGlyLys 511
QY 1392 ACGGTGAAGATTGCAAAATTTGACCGCAGCTGAGTGGTGTAGTACCAAGACAGCAGAC 1451
Db 512 ThrValAsnGluLeuGlnAsnLeuThrSerAlaGluValValProArgAspGlnThr 531
QY 1452 CCTGATCAGAACGACCAAGTTCATCGTGAATAATCATCGACATTTCTATGCCAGTCAGATG 1511
Db 532 ProAspGluAsnAspGlnValValLysIleThrGlyHisPheIleAlaSerGlnLeu 551
QY 1512 GCTCAACGGAGATCCGACATCTCTGCGCCAGGTTAAGCAGCAGCAGTCAGAACGGACAG 1571
Db 552 AlaGlnArgLysIleGlnGluIleIleSerGlnValArgGlnGlnGlnProLysPro 571
QY 1572 AGT-----AACAGGCCCGCAGCAGGAGGAG 1598
Db 572 SerAlaAlaGlyProProValAlaArgArgLys 582

RESULT 13
QY6M1
ID QY6M1 PRELIMINARY; PRT; 556 AA.
AC QY6M1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hepatocellular carcinoma autoantigen.
GN P62.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99207072; PubMed=10190901;
RA Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;
RT "A novel cytoplasmic protein with RNA-binding motifs is an autoantigen
RT in human hepatocellular carcinoma."
RL J. Exp. Med. 189:1101-1110(1999).
CC 1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF057352; AAD31596.1; -.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR GO; GO:0009386; P:translational attenuation; TAS.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR SMART; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 556 AA; 61842 MW; 1EDEFB100443DDC4 CRC64;
```

Alignment Scores:

```
Pred. No.: 5,24e-117 Length: 556
Score: 1697.00 Matches: 336
Percent Similarity: 79.32% Conservative: 63
Best Local Similarity: 66.80% Mismatches: 80
Query Match: 48.06% Indels: 24
```

```
DB: 4 Gaps: 5
US-09-270-437D-7 (1-1946) x QY6M1 (1-556)
QY 96 GCTTCCCGGGTTAGGAGCGCGAAATATCAATCCGAAATATTCACCCCGACGCTCCGATGG 155
Db 76 SerLysLysLeuArgSerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrp 95
QY 156 GAAGTACTGACAGCAGCTGCTGGCTCAGTATGTGTACGTAGAGAACTGTGACCAAGTGAAC 215
Db 96 GluValLeuAspGlyLeuLeuAlaGlnIleGlyThrValGluAsnValGluGlnValAsn 115
QY 216 ACCGAGAGTACAGCGGAGTGGTGAATGTACCTATTCACACCGGGGAGCAGACACGAGCAA 275
Db 116 ThrAspThrGluThrAlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIle 135
QY 276 GCATCATGAAGCTGAATGGCCACCTGTGAGAACCATGCCCTGAGAGGTCTCTACATC 335
Db 136 AlaMetGluLysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIle 155
QY 336 CCGGATGAGCAGATAGCA-----CAGGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGC 389
Db 156 ProAspGluGluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSer 175
QY 390 TCTCGGGGTGAGCCCGCCAGGGCTCACCTGTGCGCAGCGGGGGCCCGACCGACGAGCAG 449
Db 176 SerArgGluGlnGlyHis-----AlaProGlyGlyThrSerGlnAlaArg 190
QY 450 CAAGTGGACATCCCGCTTCGGCTCTGGTGGCCACCCAGTATGTGGGTGCCATTATTGGC 509
Db 191 GlnIleAspPheProLeuArgIleLeuValProThrGlnPheValGlyAlaIleGly 210
QY 510 AAGGAGGGGCCACCATCCGCAACATCACAACACAGACCCAGTCCAGATAGACGTGCAT 569
Db 211 LysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnSerArgValAspIleHis 230
QY 570 AGGAGGAGAACCGAGTGCAGCTGAGAAAGCCATCAGTGTGCACTCCACCCCTGAGGGC 629
Db 231 ArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleHisAlaThrProGluGly 250
QY 630 TGCTCTCCCTTGTAGATGATCTTGAGATATTCAGATATGATAAGAGGCTTAAGACACCAA 689
Db 251 ThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysGluAlaAspGluThrLys 270
QY 690 ACGGCTGACAGGTTCCCTGAAAGATCTCTGCGCCCATATAACTTTGTAGGGCGTTCATT 749
Db 271 LeuAlaGluGluIleProLeuLysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIle 290
QY 750 GGCAAGGAGGACGGAACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCATC 809
Db 291 GlyLysGluGlyArgAsnLeuLysLysIleGluHisGluThrGlyThrLysIleThrIle 310
QY 810 TCCTCGTTGCAAGACCTTACCTTTTCAACCCCTGAGAGGACCATCATCTGTGAAGGGGGCC 869
Db 311 SerSerLeuGlnAspLeuSerIleTyrAsnProGluArgThrIleThrValLysGlyThr 330
QY 870 ATCCAGAAATTTGTCAGGGCCGAGCAGGAATAATGAAGAAGTTCGGGAGGGCTATGAG 929
Db 331 ValGluAlaCysAlaSerAlaGluIleGluIleMetLysLysLeuArgGluAlaPheGlu 350
QY 930 AATGATGTGGTGGCATGAGCTCTACCTGATCCCTGGCCCTGACACCTGGCTGTGAGGT 989
Db 351 AsnAspMetLeuAlaValAsnThrHis-----Sergly 361
QY 990 CTTTTCACGACTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGTTACTGGGGCTGCT 1049
Db 362 TyrPhe-----SerSerLeuTyrProHisGln-----PheGly 373
QY 1050 CCTATAGTCTCTTTATGACAGGCTCCGAGCAGAGATGGTGCAGGTGTTATTCGCGCC 1109
Db 374 ProPheProHisHisSerTyrProGluGlnGluIleValAsnLeuPheLeProThr 393
QY 1110 CAGGCAGTGGGGCCCATCATCGCAAGAGGGGCGCAGCACATCAACAGCTCTCCCGGTTT 1169
```

01-MAY-2000 (TrEMBLrel. 13, last sequence update)
 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
 Vgl RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding protein).
 DVIIRBP.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Q., Yaniv K., Oberman F., Wolke U., Git A., Fromer M.,
 RA Taylor W., Meyer D., Standart N., Raz E., Visraeli J.K.;
 RA "Vgl RBP intracellular distribution and evolutionarily conserved
 RT expression suggest multiple roles during development.";
 RL Mech. Dev. 0:0-0(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters K.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan G.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Body;
 RX Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AF161270; RAD45610.1; -;
 DR EMBL; BC045873; AHA45873.1; -;
 DR ZFIN; ZDB-GENE-000308-1; dvirbp.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH type 1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS00084; KH TYPE 1; 4.
 DR PROSITE; PS0102; RRM; 2_1;
 SQ SEQUENCE 582 AA; 63351 MW; 9DAE63200681B306 CRC64;

Alignment Scores:
 Pred. No.: 6,96e-127 Length: 582
 Score: 1830.50 Matches: 370
 Percent Similarity: 82.19% Conservative: 50
 Best Local Similarity: 72.41% Mismatches: 74
 Query Match: 51.84% Indels: 17
 DB: 13 Gaps: 6

US-09-270-437d-7 (1-1946) x Q9PW80 (1-582)

102 CGGGTTAGGACGGGAAATTCGAATCCGAAATATTCCACCCCGAGCTCCGATGGGAAGTA 161

QY

Db 77 ArgGlnArgSerCysLysLeuGlnIleArgAsnIleProProHisMetGlnTyrGluVal 96
 QY 162 CTGCACAGCCCTGGCTCAGTATGTGTACAGTACAGAACTGTGTAGCAAGTGAACACCGAG 221
 Db 97 LeuAspGlyLeuLeuAlaGlnTyrGlyThrValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTCAGACGGCAGTGGTGAATGTACCTATTCACACCGGAGGAGCAGACCGCAGCCATC 281
 Db 117 ThrGluThrAlaValValAsnValArgTyrGlyAlaLysAspGlnAlaLysGluAlaMet 136
 QY 282 ATGAAGCTGAATGCCACCGCTGTGGAGAACCATCCCTGAAGGTCTCTATCATCCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheLeuMetGluAsnTyrAlaLeuLysValSerTyrIleProAsp 156
 QY 342 GAG---CAGATAGCACAGGACCTGAGAAATGGGGCGGAGGGGGCTTTGGCTCTCGGGGT 398
 Db 157 GluThrAlaAlaAspAlaProAlaValGlyGlyArgArgGlyPheAsnProArgGly 176
 QY 399 CAGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCCAGCAGCAGCAAGTGAC 458
 Db 177 ProProArgGlnGlySerProSerLeuGlyAlaArgPro-----LysLeuGlnSerAsp 194
 QY 459 ATCCCCCTTCGGCTCCTGGTCCCGACCCAGTATGTGGTGCATTTATGGCAAGAGGGG 518
 Db 195 ValProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGly 214
 QY 519 GCCACCATCCGCACATCACAAAACAGCAGCCAGTCCCAAGATAGAGTGCATAGGAGAG 578
 Db 215 AlaThrIleArgAsnIleThrLysGlnThrHisSerLysIleAspIleHisArgLysGlu 234
 QY 579 AACCGAGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGTGTCTCTCC 638
 Db 235 AsnAlaGlyAlaAlaGluLysProIleThrValHisSerThrProGluGlyCysSerSer 254
 QY 639 GCTGTAAAGTATCTTGGAGATATATGATATAAGAGGCTTAAGGACCAACCAACCGCTGAC 698
 Db 255 AlaCysArgAsnIleMetGluIleMetGlnLysGluAlaIleAspThrLysIleThrGlu 274
 QY 699 GAGGTTCCTCCAGAGATCCTGGCCCATTAATTAATCTTTAGGGCGTCTATTGGCAGGAA 758
 Db 275 GluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlu 294
 QY 759 GCACGAACTCTGAAGAGGTAGCAAGATACCGAGACAAATAATCACCATCTCTCTGTTG 818
 Db 295 GlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeu 314
 QY 819 CAAGACCTTTACCTTTTCAACCCCTGAGAGGACCATCATCTGTGAAGGGGCCATCGAAT 878
 Db 315 GlnAspLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyThrLeuAspAla 334
 QY 879 TGTTCAGGGCCGACGAGGAATATGAAGAAGTTCGGAGGCTTATGAGATGATGATG 938
 Db 335 CysAlaLysAlaGluGluIleMetLysLysValArgGluSerTyrGluAsnAspVal 354
 QY 939 GCTGCCATG-----AGCTCTCACCTGATCCCTGGCTGAACTGGCTCTCTAGTGTCT 992
 Db 355 AlaAlaMetHisLeuGlnSerAsnLeuIleProGlyLeuAsnLeuAlaLeuGlyLeu 374
 QY 993 TTTCCAGCTTTCATCCAGC-----GCAGTCCCGCGCTCCACG 1031
 Db 375 PheProGlyAlaAlaSerGlyGlyIleSerProSerValValSerGlyProProGly 394
 QY 1032 AGCGTTACTGGGGCTGCTCCCTTATAGTCTTATAGCTTATGAGGCTCCGAGCAGAGATGG 1091
 Db 395 AlaGlnAlaGly-----TyrGlnSerPheGlyAlaGlnMetGluSerGluThrVal 411
 QY 1092 CAGGTGTTTATCCCGCCAGGAGTGGCGCCATCATCGCAAGAAGGGGAGCAGACATC 1151
 Db 412 HisLeuPheIleProAlaLeuAlaValGlyAlaIleGlyLysGlnGlnHisIle 431
 QY 1152 AAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTGCACCCCGAAACACCTGAC 1211

OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR ENBL; AK088465; BAC40370.1; -.
DR MGD; MGI:1890359; Igf2bp3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 579 AA; 63551 MW; 937B601A95D06B77 CRC64;

Alignment Scores:

Pred. No.:	7,59e-129	Length:	579
Score:	1857.00	Matches:	377
Percent Similarity:	84.71%	Conservative:	55
Best Local Similarity:	73.92%	Mismatches:	60
Query Match:	52.59%	Indels:	18
DB:	11	Gaps:	10

US-09-270-437D-7 (1-1946) x Q8C2J9 (1-579)

QY	102	CGSGTTAGGACGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTA	161
DB	77	ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal	96
QY	162	CTGACAGCTGCTGCTCAGTAGTGACAGTGTACAGTGTGAGAACTGTGACGAAGTGAACCCGAT	221
DB	97	LeuAspSerLeuLeuValGlnIleValValGluSerCysGluGlnValAsnThrAsp	116
QY	222	ACTGAGACGCGAGTGTGTAATGTACACTATTCACACCGGAGCAGACGCGCAAGCCATC	281
DB	117	SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu	136
QY	282	ATGAGCTGAATGGCCACAGTTGGAGAACCATGCTGCTGAGTCTCTACATCCCGAT	341
DB	137	AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp	156
QY	342	GAGCAGATAGCAG-----GGACCTGAG--AATGGCGCGGAGGGGCTTTGGC	389
DB	157	GluThrAlaAlaGlnGlnAsnProSerProGlnIleuArgGlyArgGlyPro---Gly	175
QY	390	TCTGGGGTTCAGCCCGCCAGGGCTCACTGTGGCAGCGGGCGCCAGCCAGCAGCAG	449
DB	176	GlnArgGlySerSerArgGlnAlaSerPro-----GlySerValSerLysGlnLys	192
QY	450	CAAGTGGACATCCCTTCGGCTCCCTGGTCCGCCACCCAGCATATGCGGTGCGCATATTGGC	509
DB	193	ProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGly	212
QY	510	AAGGAGGGGCGCACCATCCGCAACATCACAAACAGACCCAGTCCAGATAGACGTGCAT	569
DB	213	LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis	232
QY	570	AGGAAGGAGACGAGTGCAGTGAAGAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGC	629
DB	233	ArgLysGluAsnThrGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGly	252

RESULT 12

Q9PW80

ID Q9PW80 PRELIMINARY; PRT; 582 AA.

AC Q9PW80;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

QY	630	TGCTCTCCCTGCTTAAGATGATCTTGGAGATATGCTAATAAGAGGCTAAGGACACACAAA	689
DB	253	ThrSerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGlnAspIleLys	272
QY	690	ACGGCTGACGAGTTCCTCCCTGAAGATCCTGSCCCATATAACTTTGTAGGCGCTCAT	749
DB	273	PheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle	292
QY	750	GGCAAGGAGGACGGAACCTCTGAAGAGGTAGAGCAAGATACCGGACACAAAATCACCATC	809
DB	293	GlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIle	312
QY	810	TCCTCGTTGCAAGACCTTACCTTTCAACCCCTGAGAGGACCATCATCTGTAAGGGGCC	869
DB	313	SerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlySer	332
QY	870	ATCAGAAATTTGTCAGGGCCGAGCAAAATATGAAGAAGTTCGGGAGGCGCTATCAG	929
DB	333	ValGluThrCysAlaLysAlaGluGluIleMetLysIleArgGluSerIleGlu	352
QY	930	AATGATGTGGCTGCCATGAGC-----TCTCACTGATCCCTGCGCTGAACTCGCTGCT	983
DB	353	AsnAspIleAlaSerMetAsnLeuGlnAlaAsnLeuIleProGlyLeuAsnLeuAsnAla	372
QY	984	GTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CTCCAGCAGC	1034
DB	373	LeuGlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerThr	392
QY	1035	GTTACTGGGCTGCTCCCTATAGCTCTTTATCAGGCTCCCGAGCAGGAGATGGTCAG	1094
DB	393	LeuThr-----ProProIleProGlnPheGluGlnSer---GluThrGluThrValHis	409
QY	1095	GTGTTTATCCCGCCAGGAGTGGCGCATCATCGGCAAGAGAGGGCGAGCAGCATCAAA	1154
DB	410	LeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisLys	429
QY	1155	CAGCTCTCCCGTTTCCAGCGCTCATCAAGTTCACACCCAGCAACCTGACTCTC	1214
DB	430	GlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaProAspAla	449
QY	1215	AAAGTTGCTATGTTATCATCTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGA	1274
DB	450	LysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArg	469
QY	1275	ATCTATGGCAAACTCAAGGAGGAGAACTCTTTTGGTCCCAAGGAGGAGTGAAGCTGGAG	1334
DB	470	IleTyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGlu	489
QY	1335	ACCCATACCTGTGCGCAGCATCAGCAGTGGCGGGTCTATTGGCAAGGTGGAACG	1394
DB	490	AlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThr	509
QY	1395	GTCAACGAGTTGAGAAATTTGACGCGAGCTGAGTGTGTAGTACCAAGAGACAGACCTC	1454
DB	510	ValAsnGluLeuGlnSerLeuSerAlaGluValValProArgAspGlnThrPro	529
QY	1455	GATGAGAACACACAGTCTCATCTGTAATAATCATCGACATTTCTATGCGAGTGCAGT	1514
DB	530	AspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAla	549
QY	1515	CAACGGAAGATCCGAGACATCTGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---	1568
DB	550	GlnArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeu	569
QY	1569	CAGAGTAACAGGCCCGCAGGAGGGAAG 1598	
DB	570	GlnSerGlyProProGlnSerArgArgLys 579	


```
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBU databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AK011689; BAB27779.1; -.
DR EMBL; AB046173; BAB19755.1; -.
DR EMBL; BC045138; BAH45138.1; -.
DR EMBL; BC049082; BAH49082.1; -.
DR MGD; MGI:1890359; Igf2bp3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;

Alignment Scores:
Pred. No.: 2,3e-129 Length: 579
Score: 1864.00 Matches: 378
Percent Similarity: 84.71% Conservatives: 54
Best Local Similarity: 74.12% Mismatches: 60
Query Match: 52.79% Indels: 18
DB: 11 Gaps: 10

US-09-270-437D-7 (1-1946) x Q9CPN8 (1-579)
QY 102 CGGGTAGAGCGCGGAAATTCGAATCCGAATATTCACCCAGCTCCGATGGGAAGTA 161
DB 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCGAG 221
DB 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGTGAATGTCACTATTCCACCCGGGAGCAGACCAAGCAAGCCATC 281
DB 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACCAAGTGGAGAACCAATGCCCTGAAGTCTCTACATGCCGAT 341
DB 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCACAG-----GGACCTGAG--AATGGGCGCGAGGGGGCTTTGGC 389
DB 157 GluThrAlaAlaGlnGlnAsnProSerProGlnLeuArgGlyArgArgGlyPro---Gly 175
QY 390 TCTCGGGGTGAGCCCGCCAGGCTCACCCTGTGGCAGCGGGGGCCCGCCAGCAGCAG 449
DB 176 GlnArgGlySerSerArgGlnAlaSerPro-----GlySerValSerLysGlnLys 192
QY 450 CAAAGTGACATCCCTTCGCTCTGGTGGCCACCCAGTATGTGGTGCCCATTTATGGC 509
DB 193 ProCysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGly 212
QY 510 AAGGAGGGGGCCACCATCCGCACATCACAACACAGCCAGTCCCAAGATAGACGTGCAT 569
DB 213 LysGluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHis 232
QY 570 AGAAGAGAGAACCGAGGTGAGCTGAGTGAAGAACCATCAGTGTGCATCTCCACCCCTGAGGC 629
DB 233 ArgLysGluAsnThrGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGly 252
QY 630 TGCTCTCCGCTTGTGAAGATGATCTTGGAGATATCATTAAGAGAGCTGAAGACACAA 689
DB 253 ThrSerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLys 272
```

```
QY 690 ACGGCTGACGAGGTTCCTCCGAGAGATCTCGGCCCATTAATTAATTTGTAGGGCTCTATT 749
DB 273 PheThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIle 292
QY 750 GCGAAGGAAGGACGAACTGAAGAAGGTAGACGAAGATACCGACACAAAATACCATC 809
DB 293 GlyLysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIle 312
QY 810 TCCTCGTTGCAAGACTTACCCCTTTACACCTTAGAGAGACCATCCTGTGAAGGGGCC 869
DB 313 SerProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlySer 332
QY 870 ATCGAAGATTGTTCAGGGCGGACGACGAGAAATAATGAAGAAGTTCCGGGAGCCCTATGAG 929
DB 333 ValGluThrCysAlaLysAlaGluGluIleMetLysIleArgGlnSerTyrGlu 352
QY 930 AATGATGTGGTCCCATGAGC-----TCTCACCTGATCCCTGGCTGAACCTGGCTGCT 983
DB 353 AsnAspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAla 372
QY 984 GTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CTCCACGACGC 1034
DB 373 LeuGlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerThr 392
QY 1035 GTTACTGGGGTCTCTCCCTATAGTCTTATGACAGGCTCCCGACGAGAGATGTGCAG 1094
DB 393 LeuThr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHis 409
QY 1095 GTGTTTATCCCGCCCGCAGGTCAGTGGCGCCATCATCGGCAAGAGGGGAGCACATCAAA 1154
DB 410 LeuPheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLys 429
QY 1155 CAGCTCTCCGGTTGCCAGCGCTCCATCAGATTCACACCCGACCAACACTGACTCC 1214
DB 430 GlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAla 449
QY 1215 AAGTTCGTATGTTTATCATCTGACCGCCGAGGCGCCAAATTCAGGCTCAGGGAAGA 1274
DB 450 LysValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArg 469
QY 1275 ATCTATGGCAACTCAAGGAGGAGAACTTCTTTGTTCCCAAGGAGGAAGTGAAGCTGGAG 1334
DB 470 IleTyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGlu 489
QY 1335 ACCCATACATGTCGTCAGCATCAGCAGCTGCGGGTCATTTGGCAAGAGTGGAAAAAGC 1394
DB 490 AlaHisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThr 509
QY 1395 GTGAACGAGTTCCAGAAATTTGACGCGAGCTGAGGTGGTAGTACCAAGACGACGACCCCT 1454
DB 510 ValAsnGluLeuGlnSerLeuSerSerAlaGluValValProArgAspGlnThrPro 529
QY 1455 GATGAAGACGACCGAGTCATCGTGAATAATCATCGACATTTCTATGCCAGTCAGATGGCT 1514
DB 530 AspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAla 549
QY 1515 CAAACGAGATCCGAGACATCTGCGCCCGAGTTAGCAG---CAGCATCAGAGGGA--- 1568
DB 550 GlnArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeu 569
QY 1569 CAGAGTAAACCGCCGACGCGAGGAG 1598
DB 570 GlnSerGlyProProGlnSerArgArgLys 579

RESULT 11
Q8C2J9 PRELIMINARY; PRT; 579 AA.
ID Q8C2J9
AC Q8C2J9
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGFBP3.
```

QY 513 GAGGGGCGCCATCCGCAACATCACAAACAGACCAGTCCAGATAGAGTGCATAGG 572
Db |||||
QY 214 GluGlyAlaThrIleArgAsnIleThrIleGlnSerIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGTGGAGCTGAAAGACCATAGTGTGCACCTCCACCCTGAGCGTGC 632
Db |||||
QY 234 LysGluAsnAlaGlyAlaGluGlySerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTAGATGATCTTGGAGATTATGCATAAAGAGGCTTAAGACACCAAAACG 692
Db |||||
QY 254 SerAlaAlaCysIysSerIleLeuGluIleMethIleSlyGluAlaGlnAspIleIysPhe 273
QY 693 GCTGACAGGTTCCCTGAGATCCTGGCCCATTAATACCTTTAGGGCGTCTCATGGC 752
Db |||||
QY 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGGAACCTGAAGAGTAGACAGATACCGAGACAAATAATCACCATCTCC 812
Db |||||
QY 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrIleThrIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTACCTTTACACCCCTGAGAGCCATCTACTGTGAAGGGGCCATC 872
Db |||||
QY 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATGTTGACGGCGGACGAGAAATATGAAAGTTCGGGAGGCCCTATGAGAT 932
Db |||||
QY 334 GluThrCysAlaIysAlaGluGluGluIleMetLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGCTGCATGAGC-----TCNCACTGATCCCTGCGCCGACCTGGCTGCTGTA 986
Db |||||
QY 354 AspIleAlaSerMetAsnLeuGlnAlaHisIleuLeuProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db |||||
QY 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGCTGCTGCTATAGTCTCTTATGAGCCTCCGAGCAGGAGATGTCAGGTG 1097
Db |||||
QY 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY 1098 TTTATCCCGCCAGGAGTGGCGCATCATCCGCAAGAGGAGGCGAGCACATCAACAG 1157
Db |||||
QY 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyGlyGlnHisIleGlyGln 430
QY 1158 CTCTCCCGTTTGCAGCGCTTCATCAAGATTCCACCAACCCGAAACACCTGACTCCAAA 1217
Db |||||
QY 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCTGAGCCGCGAGAGCCCAATTCAGGCTCAGGGAAGATC 1277
Db |||||
QY 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGAC 1337
Db |||||
QY 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCGAGCATCAGCAGCTGCGCGGTCTATTGCAAGGTGGAACACGGTG 1397
Db |||||
QY 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTGCAGAAATTCAGCGGAGCTGAGGTGTAGTACCAAGAGCAGACCCCTGAT 1457
Db |||||
QY 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACGAGTTCATCGTGAATAATCATCGACATTTCTATGCGAGTGCAGTGGCTCAA 1517
Db |||||
QY 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGGCCCGACGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db |||||
QY 551 ArgIysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACACGAGCCCGCAGGACGAGGAAG 1598

Db 571 SerGlyProGlnSerArgLys 579
RESULT 10 |||||
Q9CPN8 PRELIMINARY; PRT; 579 AA.
AC Q9CPN8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 10 days embryo cDNA, RIKEN full-length enriched library,
DE clone:2610035818, full insert sequence (Igf2 mRNA-binding protein 3)
DE (Insulin-like growth factor 2, binding protein 3).
GN IGFBP3 OR 2610101N1RIK OR MIMP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirul L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection.";
EL Nature 409:685-690(2001).
RN [2]
SEQUENCE FROM N.A.
RP Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
RA Yuasa Y., Takeda M., Okano H.;
RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications
RT for the developing central nervous system.";
RL J. Neurosci. Res. 0:0-0(2001).
RN [3]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Brain, and Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Toehiyuki S., Carninci P., Prange C.,
RA Raha S.S., Ioquellano N.A., Peters G.J., Abramson R.D., Mulláhy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Db 254 ThrIleHisSerThrProGluGlyCysSerAlaAlaCysLysLysIleIleMetGluIleMet 273

Qy 666 CATAAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTGAGATCTCGGCCCAT 725

Db 274 GlnLysGluAlaGlnAspThrLysPheThrGluGluIleProLeuLysIleLeuAlaHis 293

Qy 726 AATAACTTTGATGGCGCTCATTTGGCAAGGACGGAACCTGAAGGTAGACAA 785

Db 294 AsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGluGln 313

Qy 786 GATACCGAGACAAAATCACCATTCTCTGTTGCAAGACCTTACCCTTTACAACCTCGAG 845

Db 314 AspThrAspThrLysIleThrIleSerProLeuGlnAspLeuThrLeuTyAsnProGlu 333

Qy 846 AGGACCATCCTGTGAAGGGCCCATCGAATGTTTCGAGGCGCGAGCAAAATATG 905

Db 334 ArgThrIleThrValLysGlySerIleGluThrCysAlaLysAlaGluGluValMet 353

Qy 906 AGAAAGTTCGGAGGCGCTATGAGATGATGTGGCTGCCATGAGC-----TCTCACCTG 959

Db 354 LysLysIleArgGluSerTyGluAsnAspIleAlaAlaMetAsnLeuGlnAlaHisLeu 373

Qy 960 ATCCCTGGCTGAACCTGCTGCTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCG 1019

Db 374 IleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProProSerSerSerGlyMetPro 393

Qy 1020 CGCCTCCCT-----AGCAGCCTTACTGGGGTGTCCCTATAGTCTCTTTATGCGAG 1070

Db 394 ProProSerAlaGlyValSerSerProThrThrSerAlaSerTyProProPheGlyGln 413

Qy 1071 GTCCTCGAGCAGAGATGTGTGAGTGTGTTATCCCGCCGAGCAGTGGCGCCATCATC 1130

Db 414 GlnProGluSerGluThrValHisLeuPheIleProAlaLeuAlaValGlyAlaIleIle 433

Qy 1131 GCAAGAGAGGGCAGCACATCAAAAGCTCTCCCGTTTGCCAGCGCTCCATCAAGATT 1190

Db 434 GlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSerIleLysIle 453

Qy 1191 GCACCAACCGAAACCTGACTCCAAAGTTCGTATGTTATCATCTGAGCGCGCAGAG 1250

Db 454 AlaProAlaGluGlyProAspAlaLysLeuArgMetValIleIleThrGlyProProGlu 473

Qy 1251 GCCCAATTCAGGCTCAGGAGAGATCTATGCAACTGCAAGTCTCAAGGAGAGACTCTTGT 1310

Db 474 AlaGlnPheLysAlaGlnGlyArgIleTyGlyLysLeuLysGlnGluAsnPheGly 493

Qy 1311 CCCAAGGAGAGTGAAGCTGAGAGCCACATACGTGTGCCAGCATCAGCAGTGGCCGG 1370

Db 494 ProLysGluGluValLysLeuGluAlaHisIleLysValProSerTyAlaAlaGlyArg 513

Qy 1371 GTCATTGGCAAGGTGGAAAAACGGTGAACGAGTTCAGAAATTCAGCGCAGCTGAGTG 1430

Db 514 ValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGluVal 533

Qy 1431 GTAGTACCAAGAGACAGACCCCTGATGAGACGACAGTCTATCGTGAATATCATCGGA 1490

Db 534 ValValProArgAspGlnThrProAspGluAsnAspGlnValValLysIleThrGly 553

Qy 1491 CATTTCTATGCCAGTCAGATGGCTCAACGAGATCCGAGACATCTCGCCCGCAGGTTAAG 1550

Db 554 HisPheTyAlaSerGlnLeuAlaGlnArgLysIleGlnGluIleLeuAlaGlnValArg 573

Qy 1551 CAGCAGCATCAGAAG-----GGACAGAGTAAACAGGCCCGCAGCAGGAGGAG 1598

Db 574 ArgGlnGlnGlnGlnGlnLysThrAlaGlnSerGlyGlnProGlnProArgArgLys 593

RESULT 9

ID 000425 PRELIMINARY; PRT; 579 AA.

AC 000425; (T-EMBLrel. 04, Created)

DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Putative RNA binding protein KOC (KOC).

GN KOC.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;

RL Oncogene 0:0-0(0).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;

RA Mueller-Pillasch F., Lacher U., Wallrapp C.;

RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.

DR EMBL; U97188; AAC35208.1; -.

DR EMBL; U76705; AAD09223.1; -.

DR GO; GO:0005737; C:cytoplasm; TAS.

DR GO; GO:0003723; F:RNA binding; TAS.

DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR GO; GO:0006412; P:protein biosynthesis; TAS.

DR GO; GO:0006396; P:RNA processing; TAS.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH_type_1.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; rrm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS00084; KH_TYPE_1; 4.

DR PROSITE; PS0102; RRM; 2.

DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.

SQ SEQUENCE 579 AA; 63720 MW; AESC3A8EE3C135C5 CRC64;

Alignment Scores:

Pred. No.:	4,96e-130	Length:	579
Score:	1873.00	Matches:	380
Percent Similarity:	84.48%	Conservative:	50
Best Local Similarity:	74.66%	Mismatches:	63
Query Match:	53.04%	Indels:	16
DB:	4	Gaps:	9

US-09-270-437D-7 (1-1946) x 000425 (1-579)

Qy 102 CGGGTTAGAGCGCGGAAATTCAAATTCGAAATATTCACCCAGCTCGATGGAGTA 161

Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96

Qy 162 CTGGACAGCTCGCTCGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACCCGAG 221

Db 97 LeuAspSerLeuLeuValGlnTyGlyValValGluSerCysGluGlnValAsnThrAsp 116

Qy 222 AGTGACAGCGCAGTGTGAATCTCACCATTATTCACACCGGAGCAGACCGCAAGCATC 281

Db 117 SerGluThrAlaValValAsnValThrTySerSerLysAspGlnAlaArgGlnAlaLeu 136

Qy 282 ATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCCCTGAAAGTCTCTACATCCCGAT 341

Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaIleProAsp 156

Qy 342 GAGCAGATAGCA--CAGGGACCTGAGAATGGCGCCGAGGG-----GGCTTTGGCTCT 392

Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176

Qy 393 CGGGGTTCAGCCCGCCAGCGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGCAAA 452

Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193

Qy 453 GTGGACATCCCTTCGGCTCTCTGGTGGCCACCCAGTATGTGGTGGCCATTTATGGCAAG 512

Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleGlyLys 213

Db 359 SerLeuGlnSerHisLeuLeuProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1002 TCATCCAGCGAGTCCCGCCGCTCCAGCAGCGCTTACTGGGCTGCTCCCTATAGTCC 1061
Db 379 SerSerAlaValProProProProSerValThrGlyAlaAlaProTy-SerSer 398
QY 1062 TTTATGCGAGTCCCGGAGGAGATGGTGCAGGTGTTATCCCGCCAGCGAGTGGC 1121
Db 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGTGTTCAGCGCTCC 1181
Db 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTCACCCCGCAACACCTGACTCCAAAGTTCGTATGTTATCATCATCGGA 1241
Db 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleThrGly 458
QY 1242 CGGCAGAGCCCAATTCAGGGCTCAGGAGAGATCTATGCAAACTCAAGGAGGAGAC 1301
Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478
QY 1302 TTCTTTGGTCCCAAGAGGAAGTGAAGCTGAGACCCACATACCTGTGCGAGCATCAGCA 1361
Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGGCGGGTCAATGTCGCAAGTGGGAAAACCGGTGAACGAGTTCGAGAATTTGACGGCA 1421
Db 499 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGTGTAGTACCAAGAGACGAGACCCCTGATGAGACGACGAGTCACTGTAAG 1481
Db 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGGACATTTCTATGCGAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCC 1541
Db 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGGGACAGAGTAACACGAGGAGGAGGAGGAGGAG 1598
Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577

RESULT 6
O42254 PRELIMINARY; PRT; 576 AA.
AC O42254;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zipcode-binding protein.
GN ZBP1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RX MEDLINE=9722007; PubMed=9121465;
RA Ross A.F., Oleynikov Y.S., Kislauskis E.H., Taneja K.L., Singer R.H.;
RT "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL Mol. Cell. Biol. 17:2158-2165 (1997).
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF026527; AAB82295.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.

DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AAAF2D1D81C8811 CRC64;
Alignment Scores:
Pred. No.: 2,87e-168 Length: 576
Score: 2389.50 Matches: 471
Percent Similarity: 96.5% Conservative: 11
Best Local Similarity: 94.39% Mismatches: 14
Query Match: 67.67% Indels: 3
DB: 13 Gaps: 2
US-09-270-437D-7 (1-1946) x O42254 (1-576)
QY 108 AGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGAC 167
Db 79 ArgSerArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAsp 98
QY 168 AGCTGCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGA 227
Db 99 GlyLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrAspSerGlu 118
QY 228 ACGGCAGTGTGATGTACCTATTCACACCGGAGCAGACGAGCAGGAGCCATCATGAG 287
Db 119 ThrAlaValValAsnValThrTyrThrAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGCCACCAAGTTGGAGAACCATGCTGAAGGTCTCTATATCCCGATGAGCAG 347
Db 139 LeuAsnGlyHisGlnLeuGluAsnHisValSerTyrIleProAspGluGln 158
QY 348 ATAGCAGAGGACCTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGGTACGCCGCG 407
Db 159 SerValGlnGlyProGluAsnGlyArgArgGlyGlyPheGlyAlaArgGlyAlaProArg 178
QY 408 CAGGGCTCACTGTGGCAGCGGGGCGCCAGCCAGCAGCAGCAGCAGTGGACATCCCCCT 467
Db 179 GlnGlySerProValThrAlaGlyAlaProValLysGlnGlnProValAspIleProLeu 198
QY 468 CGGCTCCTGTGGCCACCCAGATATGTGGGTGCCATTATTGGCAAGGAGGGGCCACCATC 527
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 528 CGCAACATCAAAACAGACCCAGTCCAAAGTACAGTACAGTACATAGGAAGGAGAACCGAGT 587
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTCAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCTCGCTTGTAA 647
Db 239 AlaAlaGluLysAlaIleSerIleHisSerThrProGluGlyCysSerAlaAlaCysLys 258
QY 648 ATGATCTTGAGATTATGCATAAAGAGGCTTAAGCACACCAAAACCGCTGACGAGTTCCC 707
Db 259 MetIleLeuGluIleMetGlnLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTGGCCCATATAACTTTGTAGGGCGTCTCATTTGCAAGGAGGAGCGAAG 767
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAGGTAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTT 827
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrLysGlySerLeuGlnAspLeu 318
QY 828 ACCCTTTACAACCCCTGAGAGGACCATCTGTGAAGGGGGCCATCGAGATTGTTCAGG 887
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlySerIleGluAsnCysLys 338
QY 888 GCGGACGAGAAATATGAAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGTGCATG 947
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTCACCTGATCCCTGGCTGMACTGGCTGCTGTAGTCTTTTCCAGCT 1001
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378

```
QY 1242 CGCCAGAGGCCCAATTCAAGCTCAGGGAAGAACTATGGCAAACTCAGAGAGAGAAC 1301
|||
Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478
|||
QY 1302 TTCCTTTGGTCCCAAGAGGAAGCTGAAGCTGGAGAGCCACATACGTGTGCAGCATCAGCA 1361
|||
Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProLaserAla 498
|||
QY 1362 GCTGCCCGGTCATTGGCAAAAGTGGAAAAACCGTGAACAGCTTCGACGAATTTGACGGCA 1421
|||
Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGluAsnLeuThrAla 518
|||
QY 1422 CTTGAGGTGGTACTACCAAGACACGACCCCTGATGAGAACGACCAAGCTCATCTGAAA 1481
|||
Db 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
|||
QY 1482 ATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCTGGCC 1541
|||
Db 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
|||
QY 1542 CAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACCAAGCCGACGACGAGGAG 1598
|||
Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577
|||
RESULT 5
Q80US9 PRELIMINARY; PRT; 577 AA.
ID Q80US9;
AC Q80US9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Igfbp1 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051679; AAH51679.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; xrm; 2.
DR SMART; SM00322; KH; 4.
```

```
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D439F544257DA3CE CRC64;

Alignment Scores:
Pred. No.: 2.9e-175 Length: 577
Score: 2484.00 Matches: 492
Percent Similarity: 98.80% Conservative: 1
Best Local Similarity: 98.60% Mismatches: 4
Query Match: 70.35% Indels: 2
DB: 11 Gaps: 1

US-09-270-437D-7 (1-1946) x Q80US9 (1-577)
QY 108 AGAGCCCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGAGAC 167
Db 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTyrGluValLeuAsp 98
|||
QY 158 AGCCTGCTGGCTCAGTATGGTACAGTAGAGACTGTGAGCAAGTGAACACACGAGAGTGA 227
|||
Db 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
|||
QY 228 ACGGAGGTGGTCAATGTCACTATTTCACCTATTCCAAACCGGAGCAGACCAGGCAAGCCATCATGAAG 287
|||
Db 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
|||
QY 288 CTGAATGGCCACCACTGTGGAGAACCATGCCCTGAAAGTCTCTCATATCCCCGATGAGCAG 347
|||
Db 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
|||
QY 348 ATAGCACAGGACCTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGTTCAGCCCGC 407
|||
Db 159 IleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
|||
QY 408 CAGGGCTCACCTGTGGCAGCGGGGGCCCCAGGCAAGCAGCAGCAAGTGGACATCCCCCTT 467
|||
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
|||
QY 468 CGGCTCTGGTGCCCAACCCAGTATGTGGTGCCATATTGGAAGAGAGGGGGCCACCATC 527
|||
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleGlyLysGluGlyAlaThrIle 218
|||
QY 528 CGCAATCATCAAAACAGACCCAGTCCAGATAGACGTGCATAGGAGAGAGAGCAGCAGT 587
|||
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGly 238
|||
QY 588 GCAGCTGAAAAAGCCATCAGTGTGCATCTCCACCTCCAGGGCTGCTCTCCCTGCTTAAG 647
|||
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
|||
QY 648 ATGATCTTGGAGATTATGATATAAGAGGCTAAGGACACCAAAACCGCTGACAGGTTCCC 707
|||
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGlyValPro 278
|||
QY 708 CTCAGATCTCGGCCCATTAATTAATCTCTAGGGCTCTCATTTGGCAAGAGAGCAGGAC 767
|||
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
|||
QY 768 CTGAAGAGGTAGAGCAAGATACCGAGACAAAATAATCACCATCTCTCTGTTGCAAGACCTT 827
|||
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
|||
QY 828 ACCCTTTACACCCGTGAGAGGACCAATCACTGTGAAGGGGGCCCATCAGAAATTTGTCAGG 887
|||
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
|||
QY 888 GCGGAGCAGGAATAATGAGAGAACTCGGAGGCGCTATGAGAATCATGCTGCTGCCATG 947
|||
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
|||
QY 948 AGC-----TCTCACCTGATCCCTGGCTGAAACCTGGCTGCTGAGGTCTTTTCCAGCTT 1001
|||
```


RP SEQUENCE FROM N.A.
 RX MEDLINE=94179348; PubMed=81322663;
 RA Prokipcak R.D., Herrick D.J., Ross J.;
 RT "Purification and properties of a protein that binds to the C-terminal
 RT coding region of human c-myc mRNA";
 RL J. Biol. Chem. 269:9261-9269 (1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97322234; PubMed=9178888;
 RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,
 RA Gruppuso P.A., Ross J.;
 RT "Developmental regulation of CRD-BP, an RNA-binding protein that
 RT stabilizes c-myc mRNA in vitro";
 RL Oncogene 14:1279-1286 (1997).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
 RA Fleisig A.J.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustingich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Makhima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AF061563; AAC72743.1; -;
 DR EMBL; AK013940; BAB29071.1; -;
 DR HSSP; P11940; 1CVJ.
 DR MGD; MGI:1890357; Igf2bp1.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; xrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PSS0084; KH_TYPE_1; 4.
 DR PROSITE; PSS0102; RRM; 2.
 SQ SEQUENCE 577 AA; 63450 MW; EFB1AF2FF9F0344 CRC64;

 Alignment Scores:
 Pred. No.: 8,79e-176 Length: 577
 Score: 2491.00 Matches: 493
 Percent Similarity: 99.00% Conservative: 1
 Best Local Similarity: 98.80% Mismatches: 3
 Query Match: 70.55% Indels: 2
 DB: 11 Gaps: 1

 US-09-270-437d-7 (1-1946) x 088477 (1-577)

 QY 108 AGAGCCGGAATTCAGAAATTCAGAAATTCACCCAGCCTCGATGGGAAGTACTGGAC 167
 Db |||||
 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98

QY 168 AGCCTGCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGGAG 227
 Db |||||
 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
 QY 228 ACGCAGTGGTGAATGTCCACCTATTCCAAACGGGAGCAGACACGAGCAAGCCATCATGAAG 287
 Db |||||
 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetIys 138
 QY 288 CTGAATGGCCACCAAGTTCGGAACCACTCCCTGGAAGTCTCTCTACATCCCCGATGAGCAG 347
 Db |||||
 139 LeuAsnGlyHisGlnLeuLeuAsnHisAlaLeuIysValSerTyrIleProAspGluGln 158
 QY 348 ATAGCACAGGACCTGACAATGGGGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCC 407
 Db |||||
 159 IleThrGlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
 QY 408 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCCCTT 467
 Db |||||
 179 GlnGlySerProValAlaAlaGlyAlaProAlaIysGlnGlnProValAspIleProLeu 198
 QY 468 CGGCTCCTGGTCCCAACCCAGTATGTGGTGCCATTATTGGCAAGAGGGGCCACCATC 527
 Db |||||
 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
 QY 528 CGCAACATCACAACACACACCCAGTCCCAAGATAGACGTGCATAGNAGAGAGACCCAGGT 587
 Db |||||
 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
 QY 588 GCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGTCTCTCCGCTGTGAAG 647
 Db |||||
 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
 QY 648 ATGATCTTGGAGATTATCATTAAGAGGCTAAGGACACCAAAACCGCTGACAGAGTTCC 707
 Db |||||
 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
 QY 708 CTGAAGATCTGGCCCAATAAATCTTCTAGGGCTCTCATTTGGCAAGGAGGCGGAC 767
 Db |||||
 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
 QY 768 CTGAAGAGGTAGACCAAGATACCCAGACAAAATCACCATCTCTCTGTTGCAAGACCTT 827
 Db |||||
 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
 QY 828 ACCCTTTACAAACCTGAGAGGACCATCATCTGTAAGGGGGCCATCGAGAAATTTGTGAGG 887
 Db |||||
 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
 QY 888 GCCGAGCAGGAATAATGAAGAAAGTTCGGGAGGCTATGAGAATGATGTGCTGCCATG 947
 Db |||||
 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
 QY 948 AGC-----TCTCACCTGATCCCTGGCTGAAACCTGGCTGTGTAGTCTTTTCCAGCT 1001
 Db |||||
 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
 QY 1002 TCATCCAGGAGTCCCGCGCTCCAGCGAGTTCCTGGGGCTGCTCCCTATAGTCTCC 1061
 Db |||||
 379 SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTyrSerSer 398
 QY 1062 TTTATGCAAGCTCCCGAGCAGAGATGTGCAGGTGTTTATCCCGCCCGCAGGAGTGGGC 1121
 Db |||||
 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
 QY 1122 GCCATCATCGCAAGAGGGCAGCACATCAACACAGCTCTCCCGGTTTGGCCAGCGCTCC 1181
 Db |||||
 419 AlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
 QY 1182 ATCAAGATTGCACCCCGAACAACCTGACTCCAAAGTTCTGATGTTATCATCATCTGGA 1241
 Db |||||
 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValValIleThrGly 458

DR Pfam; PF00076; rim; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE 1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;

Alignment Scores:
Pred. No.: 2,25e-176 Length: 577
Score: 2499.00 Matches: 495
Percent Similarity: 99.20% Conservative: 0
Best Local Similarity: 99.20% Mismatches: 2
Query Match: 70.77% Indels: 2
DB: 4 Gaps: 1

US-09-270-437d-7 (1-1946) x Q9NZI8 (1-577)

QY 108 AGGAGCGGAAATTCAGAAATTCACCCAGCTCCGATGGAGTACTGAC 167
Db 79 ArgSerArgIleGlnIleArgAsnIleProProGlnIleuArgTrpGluValLeuAsp 98

QY 168 AGCTGTCTGCTCAGTGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGA 227
Db 99 SerLeuLeuAlaGlnTyrglyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118

QY 228 ACGGAGTGTGATGTCACTATTCCAAACCGGAGCAGACCAGGCAAGCCATCATGAAG 287
Db 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138

QY 288 CTGAATGGCCACAGTGGAGAACCATCCCTGAAGTCTCCTACATCCCGATGAGCAG 347
Db 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerYrIleProAspGluGln 158

QY 348 ATAGCAGAGGACCTCAGAAATGGCGCGGAGGGGCTTTGGCTCTCGGGGTTCAGCCCGC 407
Db 159 IleAlaGlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178

QY 408 CAGGCTCAGTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAGCAGTGGACATCCCCCTT 467
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198

QY 468 CGGCTCCTGTGCGCCACCATGTGTGGTGGCCATTATGGCAAGGAGGGGCCACCATC 527
Db 199 ArgLeuLeuValProThrGlnTyrglyValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218

QY 528 CGCAACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGAGAGAGAGACGACGAGT 587
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238

QY 588 GCAGCTGAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCTCGCTGTGAAG 647
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258

QY 648 ATGATCTTGAGATTATGCAAGAGCGCTAAGACACCAAAACGGCTGACGAGTTCCC 707
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278

QY 708 CTGAAGATCCTGGCCCAATAACTTTGTAGGGCGTCTCATTTGGCAAGGAGGACGGAAC 767
Db 279 LeuLysThrLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298

QY 768 CTGAGAGGTGAGAGCAAGATACCGAGCAAAATCAATCTCCTCTGTTGCAAGACCTT 827
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318

QY 828 ACCCTTTACACCTGAGAGACCATCATCTGTGAAGGGGGCCATCGAGAAATTTGTGAGG 887
Db 319 ThrLeuTyrAsnProGluuArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338

QY 888 GCCGAGCAGGAAATAATCAAGAAAGTTCCGAGGCGCTATGAGATGATGTGGCTGCCATG 947
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358

QY 948 AGC-----TCTCACCTGATCCCTGGCTGAACCTGGCTGCTAGGTCCTTTCCAGCT 1001
Db 359 SerLeuGlnSerHisLeuThrProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378

QY 1002 TCATCCAGCGCAGTCCCGCCCTCCAGAGCGCTTACTGGGCTGTCTCCCTATAGCTCC 1061
Db 379 SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTyrSerSer 398

QY 1062 TTTATGAGGCTCCCGAGCAGGAGATGTCAGGTGTTTATCCCGCCCGAGGAGTGGC 1121
Db 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418

QY 1122 GCATCATCGCGAAGAGGGCAGCAGCATCAACAGCTCTCCCGGTTGGCAGGCGCTCC 1181
Db 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438

QY 1182 ATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCACTGGA 1241
Db 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleIleThrGly 458

QY 1242 CCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGAGAAC 1301
Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478

QY 1302 TTTCTTGGTCCCAAGAGGAGTCAAGCTCGAGAGCCACATACCTGTGCCAGCATCAGCA 1361
Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498

QY 1362 GCTGGCCGGTCTATTGGCAAGGTGGAACACGGTGAACGAGTTGCAAGATTTACGGCA 1421
Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518

QY 1422 GCTGAGTGTGTAGTACCAAGAGACCAACCCCTGATGAGACCAACGAGTTCATCGTAAA 1481
Db 519 AlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538

QY 1482 ATCATCGACATTTCTATGTCAGTTCAGTTCAGTCAAGCGAGATCCGAGACATCCTGCC 1541
Db 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558

QY 1542 CAGTTTAAGCAGCAGCATCAGAGGGACAGATTAACCGGCCAGGACGAGGAGAG 1598
Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 577

RESULT 3

ID 088477 PRELIMINARY; PRT; 577 AA.
AC 088477;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coding region determinant binding protein.
GN IGF2BP1 OR CRDBP.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=92217743; PubMed=1559612;
RX Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;
RA "Control of c-myc mRNA half-life in vitro by a protein capable of
RT binding to a coding region stability determinant.";
RL Genes Dev. 6:642-654(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94156886; PubMed=8114742;
RA Herrick D.J.; Ross J.;
RT "The half-life of c-myc mRNA in growing and serum-stimulated cells:
RT influence of the coding and 3' untranslated regions and role of
RL ribosome translocation.";
RL Mol. Cell. Biol. 14:2119-2128(1994).
RN [3]

DR Pfam; PF00076; rtm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBDLEE CRC64;

Alignment Scores:
Pred. No.: 9.59e-177 Length: 577
Score: 2504.00 Matches: 495
Percent Similarity: 99.40% Conservative: 1
Best Local Similarity: 99.20% Mismatches: 1
Query Match: 70.91% Indels: 2
DB: 11 Gaps: 1

US-09-270-437D-7 (1-1946) x Q8CGX0 (1-577)

QY 108 AGGAGCCGGAATAATTCAATCCGAATAATTCCACCCAGCTCCGATGGGAAGTACTGGAC 167
DB 79 ArgSerArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAsp 98
QY 168 AGCTGCTGGCTCAGTATGAGTACAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
DB 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
QY 228 AGCGAGTGGTCAATGCTCACTATTCCACCGGGAGCAGACCGAGCCATCATGAG 287
DB 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGGCCACAGTGGAGAACCATGCCCCCTGGAAGTCTCTACATCCCCGATGAGCAG 347
DB 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
QY 348 ATAGCACAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGTCAGCCCCGC 407
DB 159 IleAlaGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGCTCACTGTGCGACGGGGCCCGAGCCAGCCAGCAGCAGTGGACATCCCCCTT 467
DB 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
QY 468 CGGCTCTCGTCCCCACCCAGATGTGGGTGGCTATTATTGGCAAGAGGGGGCCACCATC 527
DB 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluAlaThrIle 218
QY 528 CGCAATCATCAAAACAGACCCAGTCCCAAGATAGAGTGCATAGAGAGAGACCGAGT 587
DB 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTGAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGTCTCTCCGCTTGAAG 647
DB 239 AlaAlaGluLysAlaLysSerValHisSerThrProGluGlyCysSerAlaCysLys 258
QY 648 ATGATCTTGGAGATTATGCATAAAGAGGTGAAGCACCAAAACCGCTGACGAGTTCCC 707
DB 259 MetIleLeuGluIleMethisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTGGCCCATATATCTTTGTAGGCGTCTCATTTGGCAAGGAAGACCGGAC 767
DB 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAAGGTAGCAAGATACCGAGACCAAAATCACCATCTCTCTGTTGCAAGACCTT 827
DB 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTTACACCCCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTTCAGG 887
DB 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
QY 888 GCGCAGAGAAATATAGAAAGTTCCGGAGGCTTATGAGAAATCATGTGGCTGCGCATG 947
DB 339 AlaGluGlnIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358

QY 948 AGC-----TCTCACTGATCCCTGGCCTGAACCTGGCTGCTGAGGCTTTTCCAGCT 1001
DB 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1002 TCATCCAGCGCAGTCCCGCCGCTCCCAAGCAGCGTTACTGGGGTCTCCCTTATAGCTCC 1061
DB 379 SerSerSerAlaValProProProSerValThrGlyAlaAlaProTyrGlySer 398
QY 1062 TTTATGAGGCTCCCGAGCAGAGATGTGTGAGGTGTTTATCTCCCGCCAGCAGTGGGC 1121
DB 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGCAAGAGGGCAGCACATCAAAAGCTCTCCCGTTGGCCAGCGCTCC 1181
DB 419 AlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGACCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCTGGA 1241
DB 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleThrGly 458
QY 1242 CGCCAGAGGCCCAATTCAAGCTCAGGAGAGAACTCTATGGCAAACTCAAGAGGAGAC 1301
DB 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluAsn 478
QY 1302 TTCTTTGTTCCTCAAGCAGAGAGTGAAGCTGAGAGACCCACATACGTGTGCCATCAGCA 1361
DB 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGCGCGGCTCATTTGGCAAGGTGGAACCGTGAACGAGTTCAGAAATTTGACGGA 1421
DB 499 AlaGlyArgValIleGlyLysGlyThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGGTGCTAGTACCAAGACAGACCCCTGTATGAGACGACGCTCATCTGGA 1481
DB 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGGACATTTCTATCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGGCC 1541
DB 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGAGGACAGATAACAGGCCCGCCAGGAGGAG 1598
DB 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 577

RESULT 2
Q9NZ18 PRELIMINARY; PRT; 577 AA.
AC Q9NZ18; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE mRNA-binding protein CRDBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,
RA Panoutsakopoulos G., Kyriazoglou I., Voutzoulis S., Tsiapalis C.M.,
RA Kittas C., Agnantis N., Pandis N.;
RT "Ectopic expression of a KH-domain containing protein, highly
RT homologous to both human IMP-1 and mouse CRD-BP, in benign and
RT malignant mesenchymal tumors.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF198254; AAF37203.1; -.
DR HSSP; PI1940; 1CVJ.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:55:00 ; Search time 93.0388 Seconds
(without alignments)
13198.757 Million cell updates/sec

Title: US-09-270-437D-7

Perfect score: 2531

Sequence: 1 gcctagcgaggggctggg.....attcttcagggttttaaaa 1946

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_23072004_095739_27425/app_query.fasta_1.5582
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1 1 287 @runat_23072004_095739_27425 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	70.9	577	11	Q8CGX0 rattus norv

ALIGNMENTS

RESULT 1

Q8CGX0 ID Q8CGX0 PRELIMINARY; PRT; 577 AA.

AC Q8CGX0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B-actin zipcode binding protein 1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

[1]

SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley;

RA Eom T., Singer R.H., Bassell G.J.;

RT "Molecular interactions between rzbP1 and b-actin zipcode required for
transport of mRNA and stimulation of spine growth.";

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF541940; AA016210.1;

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH type 1.

DR InterPro; IPR000504; RNA_rec_mot.

PFam; PF00013; KH; 4.

Q9NZI8 homo sapien
Q88477 mus musculu
Q8BRH1 mus musculu
Q80US9 mus musculu
Q42254 gallus gall
Q73932 xenopus lae
Q57526 xenopus lae
Q00425 homo sapien
Q9CPN8 mus musculu
Q8C2J9 mus musculu
Q9PW80 brachydanio
Q9V6M1 homo sapien
Q7TF99 mus musculu
Q9VZ69 drosophila
Q9V699 drosophila
Q8IGK4 drosophila
Q21605 caenorhabdi
Q9C553 arabidopsis
Q7TP50 rattus norv
Q74919 schizosacch
Q84MA6 arabidopsis
Q9XI71 arabidopsis
Q23487 caenorhabdi
Q7Z145 caenorhabdi
Q7ZVK5 brachydanio
Q8AX85 xenopus lae
P91393 caenorhabdi
Q8S7G1 oryza sativ
Q8LNT9 oryza sativ
Q8BLA0 caenorhabdi
Q7ZXS1 xenopus lae
Q9SR13 arabidopsis
Q9LXF5 arabidopsis
Q86EP6 homo sapien
Q82762 arabidopsis
Q17935 caenorhabdi
Q17936 caenorhabdi
Q9ASX3 arabidopsis
Q9FNK3 arabidopsis
Q9L128 oryza sativ
Q8LIDV1 arabidopsis
Q8AX84 xenopus lae
Q95829 drosophila
Q9W702 xenopus lae

This Page Blank (uspto)

```
Db 418 nGluileArgLysArgThrAsnAlaAsnIleCysIleSerLysGlyLysLysAspAspLe 438
QY 827 TACCCCTTTACACCCCTGAGAGACATCAGTGTGAAGGGGCCCATCGAGAATTGTGCAG 886
Db 438 u-----ValGluValSerGlyGluValSerSerValArgAs 450
QY 887 GGCCGAGCAGGAATAATGAAGAAGTTCGGAGGCCCTATGAGAAATGATGTGGCTGCCAT 946
Db 450 pAlaLeuIleGlnIleValLeuArgLeuArgGluAspValLeuGlyAspLysAspSerVa 470
QY 947 GAGCTCTCACCTGATCCCT----- 965
Db 470 lAlaThrArgLysProProAlaArgThrAspAsnCysSerPheLeuSerGlySerSerAs 490
QY 966 ----GGCCTGAACCTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGCGAGTCCCG-- 1019
Db 490 nAlaGlyTyrThrLeuProSerPheMetSerSerMetAlaSerThrSerGlyPheHisGl 510
QY 1020 -----CGCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTAT 1066
Db 510 yTyrGlySerPheProAlaGlyAspAsnValLeuGlySerThrGlyProTyrSerTyrGl 530
QY 1067 GCAGGCTCCCGAGCAGGAGATGGTGCAAGTGTGTTATCCCGCCCGCAGTGGCGCCAT 1126
Db 530 yArgLeuProSerSerSerAlaLeuGluLeuLeuProAlaHisAlaMetSerLysVa 550
QY 1127 CATCGCAAGAAGGGCGACACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAA 1186
Db 550 lMetGlyLysGlyGlyGlyAsnLeuGluAsnIleArgArgIleSerGlyAlaMetIleGl 570
QY 1187 GATTGCACCCCGAACA 1205
Db 570 uIleSerAlaSerLysThr 576
```

Search completed: July 23, 2004, 11:06:27
Job time : 41.1191 secs

RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirion A., Vitale D., Liguri R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller K., Berger C., Monfort A., Casacuberta E.,
RA Chefdor F., Cooke R., Vandenbol M., Barges M., Terol J., Torres A.,
RA Gibbons T., Weber N., Vandenbol M., Barges M., Tacon D., Jesse T.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Francis P., Biele C.,
RA Heijnen L., Schwarz S., Scholler P., Heber H.-W., Stocker S.,
RA Prishan D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekon M., Murray J., Sheet P., Cordes M., Abuthreiden J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.W.,
RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M.A., Martienssen R., McCombie W.R.; the
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";
RT Nature 402:769-777(1999).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J.M., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.F., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kaniya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamez R., Vayberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846(2003).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P58223-1; Sequences=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=P58223-2; Sequences=VSP_008899, VSP_008900;
CC Note=May be due to a competing acceptor site. No experimental
CC confirmation available;
CC -!- SIMILARITY: Contains 5 KH domains.
CC -!- CAUTION: Ref.1 sequences differ from that shown due to erroneous
CC gene model prediction. At4g18370 and At4g18375 were originally
CC fused into a single gene.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL021710; CAA16717.1; ALT_SEQ.
CC EMBL; AL161548; CAB78839.1; ALT_SEQ.
CC EMBL; AY133701; AAM91635.1; -.
CC EMBL; BT001108; AAN64172.1; -.
CC InterPro; IPR004087; KH_dom.

DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 5.
DR SMART; SM00322; KH; 5.
DR PROSITE; PS00084; KH TYPE 1; 5.
KW Hypothetical protein; Nuclear protein; RNA-binding; Repeat;
KW Alternative splicing;
FT DOMAIN 99 KH 1.
FT DOMAIN 138 210 KH 2.
FT DOMAIN 311 380 KH 3.
FT DOMAIN 394 455 KH 4.
FT DOMAIN 535 599 KH 5.
FT VARSPLIC 532 532 L -> F (in isoform 2).
FT VARSPLIC 533 606 Missing (in isoform 2).
FT VARSPLIC 533 606 /FTid=VSP_008899.
SQ SEQUENCE 606 AA; 65760 MW; 61F135BBB8647C0C CRC64;
Alignment Scores: Length: 606
Pred. No.: 3.98e-06 Matches: 89
Score: 216.50 Conservative: 80
Percent Similarity: 43.67% Mismatches: 164
Best Local Similarity: 23.00% Indels: 54
Query Match: 6.13% Gaps: 10
DB: 1
US-09-270-437d-7 (1-1946) x Y475_ARATH (1-606)
QY 125 AATCGAAATATTCACCCAGCTCCGATGGGAAGTACTGACAGCGCTGGTGGTCACTA 184
DB 218 AsnProArgGluAsnIleProLeuAspSerThrSerGlnAspValProAlaAspVal 237
QY 185 TGTACACTAGAGAACTGTGACAGTCAACACCGAGAGTGAAGCGGCGGCGGCGG 238
DB 238 IleValProSerAspLeuSerAsnSerValTyProGln---ThrGlyPheTySerAsn 256
QY 239 GAATGTCACCTATTCCACCGGAGCAGACACCGAGGAGCATCATGAAGCTGAATGGCA 298
DB 257 GlnAspHisIleLeuGlnGlnGlyAlaGlyValProSerTyPheAsnAla----- 273
QY 299 CCAGTTGGAGAACCATGCCCTGGAAGGTCTCTACATCCCGAGTACGAGCAGACAGG 358
DB 274 -----LeuSerValSerAspPheGlnGly 281
QY 359 ACCTGAGATGGCGCGCGGGGGTGGTCTCGGGTTCAGCCCGCGGCTCAC 418
DB 282 TyrAlaGluThrAlaAlaAsnProValProValPheAlaSerSerLeuPro---ValThr 300
QY 419 TGTGCGAGCGGGCGCGCGCGAGCAGCAGCAAGTGGACATCCCTTCGGCTCTGCTGT 478
DB 301 HisGlyPheGlyGlySerSer-ArgSerGlu-----GluLeuValPheLysValLeuCy 318
QY 479 GCCACCCAGTATGTGGTGGCGCATTTATGGCAAGAGGGGGCGCCACCATCCGCAACATC 538
DB 318 sProLeuCyAsnIleMetArgValIleGlyLysGlySerThrIleLysArgIle 338
QY 539 AAAACAGACCCAGTCCACAGTACAGCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGG 598
DB 338 gGluAlaSerGlySerCysIleGluValAsnAspSerArgThrLysCysGlyAspAspG 358
QY 599 AGCCATCTAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCCTGCTGTGTAAGATGATCTTGA 658
DB 358 uCysValIleIleValThrAlaThrGluSerProAspMetLysSerMetAlaValG 378
QY 659 G-----ATTATGCAATAAGAGCTTAAGACACCAACCGCTGACGAGGTTCCTCGTAA 712
DB 378 uAlaValLeuLeuGlnGluTyriLeAsnAspGluAspAlaGluAsnValLysMetG 398
QY 713 GATCTGGGCCATAATACTTTGTAGGGCGTCTCATTTGGCAGGAGGAGGAGGAGGAGGAG 772
DB 398 nLeuLeuValSerSerLysValIleGlyCysValIleGlyLysSerGlySerValIleAs 418
QY 773 GAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC-----TCGTTGCAAGACCT 826
DB 773 GAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC-----TCGTTGCAAGACCT 826

443	AspHisLysPheHisArgHisLeuIleGlyLeuSerGlyAlaAsnIleAsnArgIleLys	462
339	GATGAG-----CAGATAGCACAGGAGCTGAGAAATGGGCGCCGAGGGGCTTTGGC	389
463	AspGlnTyrLysValSerValArgIleProProAspSerGlnLysSerAsnLeuIleArg	482
390	TCTCGGGGTGAGCCGCCAGGGGTCACTGTGGCGAGCGGGGGCCCCAGCCAGCAGCAG	449
483	IleGluGlyAspPro---GlnGlyVal-----GlnGlnAlaLysArgGlu	496
450	CAAGTGACATCCCCCTTCGG-----CTCTGGTGGCCCAACC	485
497	LeuLeuGluLeuAlaSerArgMetGluAsnGluArgThrLysAspLeuIleIleGluGln	516
486	CAGTATGTGGGTGCATTATTGGCAAGGAGGGGCCACCATCGCAACATCACAAACAG	545
517	ArgPheHisArgThrIleIleGlyGlnLysGlyGluArgIleArgGluIleArgAspLys	536
546	ACCAGTCCCAAGATA-----GACGTGCATAGGAAGAGAACGCGGTGCAGCT	593
537	PheProGluValIleIleAsnPheProAspProAlaGlnLysSerAsp-----	552
594	GAAAAAGCCATCAGTGTGCACTCACCCCTGAGGGGTGCTCCTCCGCTTGTAGATGATC	653
553	-----IleValGlnLeuArgGlyProLysAsnGluValGlnLysCysThrLysTyr	569
654	TTGAGATTATGCATAAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCTCCCT	713
570	MetGlnLysMetValAlaAspLeuValGluAsnSerTyrSerIleSerValProIlePhe	589
714	ATCTGCGCCCAATAAATCTTTGTAGGCGCTCATTTGGCAAGGAAGACCGAACTGAAG	773
590	LysGlnPheHisLysAsn-----IleIleGlyLysGlyAlaAsnIleLys	605
774	AAGTAGAGCAAGATACCGACAAAAATCACCATCTCCTCGTTGCAAGACCTTACCCCT	833
606	LysIleArgGluGluSerAsnThrLysIle-----AspLeuProAla	619
834	TACAACCT--GAGAGACCATCACTGTGAAGGGGGCCATCGAGAATTTGTGCGAGGCC	890
620	GluAsnSerAsnSerGluThrIleIleIleThrGlyLysArgAlaAsnCysGluAlaAla	639
891	GAGCAGAAATAATCAAGAAAGTTTCGGAGGCGCTATGAGAATGATGTGGCTCCATGAGC	950
640	ArgSerArgIleLeu-----SerIleGlnLysAspLeuAla-----	651
951	TCTCACCTGATCCCTGGCTGAACTGGCTGCTAGTGTCTTTTCCAGCTTCATCCAGC	1010
652	-----AsnIleAlaGlu-----	655
1011	GCAGTCCCGCGCCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGTCTCTTTATGCGAG	1070
655	-----	655
1071	GCTCCCGACGAGATGGTGCAGGTGTTATPCCCCGCCAGCAGGTGGCGCGCATCATC	1130
656	-----ValGluValSerIleProAlaLysLeuHisAsnSerLeuIle	669
1131	GGCAAGAGGGGCAGCAGCATCAACAGCTCTCCGGTTTGGCCAGCGGCTCCATCAAGATT	1190
670	GlyThrLysGlyArgLeuIleArgSerIleMetGluGluCysGlyGlyValHisIleHis	689
1191	GCACCACCGGAAACCACTGACTCCAAAGTTCGTATGTTTATCATCTGACCGCCAGAG	1250
690	PheProValGluGlySerGlySerAspThr-----ValValIleArgGlyProSerSer	707
1251	GCCCAATTCAAGGCTCAGGGGAAGAAATCTATGCAAACTCAAGGAGGAGAACTCTTTGGT	1310
708	AspValGluLysAlaLysLysGlnLeuLeu---HisLeuAlaGluGlu-----	722
1311	CCCAAGGAGGAAGTGAAGCTGGAGCCCAATACATACGTGTGGCAGCATCAGAGCTGGCGG	1370

723 ---LysGlnThrLysSerPheThrValAspIleArgAlaLysProGluTyrHisLysPhe 741

1371 GTCATTGGCAAAAGGTGGAAAAACGGTGAACAGAGTTCGACAGATTTCAGCGCAGCTGAGGTG 1430

742 LeuIleGlyIysGlyGlyGlyLysIleArgLysValArgAspSerThrGlyAlaArgVal 761

1431 GTAGTACCAAGACAGACACCCCTGATGAGAACGACGACAGGTATCTGTGAAAATCATCGGA 1490

762 IlePhePro-----AlaAlaGluAspLysAspGlnAspLeuIleThrIleGly 778

1491 CATTTCTCTAGTCAGATGGCTCAACGGAACATCCGAGACATCTCTGGCCCGAGTTAAG 1550

779 LysGluAspAlaValArgGluAlaGlnLysGluLeuGluAlaLeuIle-GlnAsnLeuAs 798

1551 CAGCAGCATCATCAAGGAGGACAGAGTAACGAGGCCAGGCAC 1590

798 pAsnValValGluAspSerMetLeuValAspProLysHis 811

RESULT 14

PCB4_HUMAN	STANDARD;	PRT;	403 AA.
ID	PCB4_HUMAN		
AC	P57723;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Poly(rC)-binding protein 4 (Alpba-CP4).		
GN	PCBP4.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20396135; PubMed=10936052;		
RA	Makeyev A.V., Liehaber S.A.;		
RT	"Identification of two novel mammalian genes establishes a subfamily		
RT	of KH-domain RNA-binding proteins.";		
RL	Genomics 67:301-316 (2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,		
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,		
RA	Wagatsuma M., Hosoi T., Kaku Y., Kondo H., Sugawara M.,		
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,		
RA	Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,		
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,		
RA	Ninomiya K., Iwayanagi T.;		
RT	"NEDO human cDNA sequencing project.";		
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Lung;		
RX	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahay J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bonifard G.G.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,		
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length		
RT	human and mouse cDNA sequences.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		

Db 238 LeuAsnGlnValAlaArgGlnGlnSerHisPheAlaMetMetHisGlyGlyThrGlyPhe 257
QY 1029 AGCAGCGTACTGGGGTGTCTCC-----TATAGTCTCTTTATGACGCTCCC 1076
Db 258 AlaglyIleAspSerSerProGluValIleGlyGlyTrpAlaSerLeuAspAlaSer 277
QY 1077 GAGCAGAGATGGTGCAGGTGTATATCCCGCCAGCAGTGGGGCCATCATCGGCAAG 1136
Db 278 ThrGlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleIleGlyArg 297
QY 1137 AAGGGGCGACACATCAACAGCTCTCCGGTTCGCCAGCTCCATCAAGATGACCA 1196
Db 298 GlnGlyAlaAsnIleAsnGluIleArgGlnMetSerGlyAlaGlnIleIleValAlaAsn 317
QY 1197 CCCGAAACACTGACTCAAAAGTTTCGTATGTTATCATCTGACCGCCAGCGCCAA 1256
Db 318 Pro---ValGluGlySerSerGlyArgGlnValThrIleThrGlySerAlaAlaSerIle 336
QY 1257 TTCAGGCTCAGGAGAACTATGCGAACTCAAGGAGGAGAACTCTTTGGT 1310
Db 337 SerIleuAlaGlnIleValLeuIleAsnAlaArgLeuSerSerGluIleGlyMetGly 354

RESULT 13

VGLN_HUMAN
ID VGLN_HUMAN STANDARD; PRT; 1268 AA.
AC Q00341; Q9UCY3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Viginin (High density lipoprotein-binding protein) (HDL-binding
DE protein).
GN HDLBP OR HBP OR VGL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92291094; PubMed=1318310;
RA McKnight G.L., Reasoner J., Gilbert T., Sundquist K.O., Hokland B.,
RA McKernan P.A., Champagne J., Johnson C.J., Bailey M.C., Holly R.,
RA O'Hara P.J., Oram J.F.,
RT "Cloning and expression of a cellular high density lipoprotein-binding
RT protein that is up-regulated by cholesterol loading of cells.";
RL J. Biol. Chem. 267:12131-12141(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohyiuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SUBCELLULAR LOCATION.
RX MEDLINE=96184515; PubMed=8605996;
RA Kugler S., Grunweller A., Probst C., Klinger M., Muller P.K.,

RA Kruse C.;
RT "Viginin contains a functional nuclear localisation sequence and is
RT present in both the cytoplasm and the nucleus.";
RL FEBS Lett. 382:330-334(1996).
CC -!- FUNCTION: APPEARS TO PLAY A ROLE IN CELL STEROL METABOLISM.
CC IT MAY FUNCTION TO PROTECT CELLS FROM OVER-ACCUMULATION OF
CC CHOLESTEROL.
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -!- SIMILARITY: Contains 14 KH domains.
CC -!- SIMILARITY: STRONG, TO C.ELEGANS C08H9.2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; M64098; AAA35962.1; -.
DR EMBL; BC001179; AAH01179.1; -.
DR PIR; A44125; A44125.
DR PDB; 1VIG; 03-APR-96.
DR PDB; 1VIH; 03-APR-96.
DR Genew; HGNC:4857; HDLBP.
DR MIM; 142695; -.
DR GO; GO:0005886; C:plasma membrane; TAS.
DR GO; GO:0008289; F:lipid binding; TAS.
DR GO; GO:0008203; P:cholesterol metabolism; TAS.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 14.
DR SMART; SM00322; KH; 12.
DR PROSITE; PS00084; KH type 1; 14.
KW HDL; Lipid transport. Cholesterol metabolism; RNA-binding; Repeat;
KW Phosphorylation; Nuclear protein; 3D-structure.
FT DOMAIN 158 229 KH 1.
FT DOMAIN 230 302 KH 2.
FT DOMAIN 303 371 KH 3.
FT DOMAIN 372 442 KH 4.
FT DOMAIN 443 514 KH 5.
FT DOMAIN 515 588 KH 6.
FT DOMAIN 589 660 KH 7.
FT DOMAIN 661 734 KH 8.
FT DOMAIN 735 807 KH 9.
FT DOMAIN 808 880 KH 10.
FT DOMAIN 881 979 KH 11.
FT DOMAIN 980 1059 KH 12.
FT DOMAIN 1060 1134 KH 13.
FT DOMAIN 1135 1209 KH 14.
FT MOD_RES 295 295 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 296 296 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 1268 AA; 141439 MW; 374ECB83D13A7431 CRC64;
Alignment Scores:
Pred. No.: 3,85e-06 Length: 1268
Score: 218.00 Matches: 109
Percent Similarity: 38.46% Conservative: 81
Best Local Similarity: 22.06% Mismatches: 191
Query Match: 6.17% Indels: 113
DB: 1 Gaps: 19
US-09-270-437d-7 (1-1946) x VGLN_HUMAN (1-1268)
QY 195 GAGAACTGTGAGCAAGTGAACCGAGAGTGAGACGCGAGTGTGATGTCACCTATTC 254
Db 403 GlnGlyGluAspGlySerIleThrLeuGluGlyProThrGluAspValAsnValAlaGlnGlu 422
QY 255 AACCGGCGAGCAG-----ACCAGGCAAGCCCATCATGAAGCTGAAT--- 293
Db 423 GlnIleGluGlyMetValIleAsnArgMetAspTyrValGluIleAsnIle 442
QY 294 GGCACACAGTTGGAGAACCATGCCCTG-----AAGTCTCTTACATCCCC 338

QY	998	-----	999	-----
Db	290	LeuSerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAla	309	
QY	999	-----	1010	-----GCTTCATCCAGC
Db	310	SerTyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuSerGlnAlaAlaThrGly	329	
QY	1011	CGAGTCCCGCGCTCCAGCAGCGTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAG	1070	
Db	330	AlaLeuAlaAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaLeuLeuAla	349	
QY	1071	GCTCCGAGCAGCAGATGGTCAGGTGTATTCCCGCCAGGCGAGTGGCGCCATCATC	1130	
Db	350	ThrTyrAlaSerGlu-----	362	-----AlaSerAlaSerGlySerThrAla
QY	1131	GGCAAGAGGGGAGCAGCATCAAACAGCTCTCCGGTTTCGCCAGCGCTCCATCAAGATT	1190	
Db	363	GlyGlyThrAla-----	376	-----GlyThrPheAlaLeuGlySerLeuAlaAla
QY	1191	GCACACCCGAAACACCTGACTCAAAGTTCGATGTTATCATCTGACCGCCAGAG	1250	
Db	377	Ala-----	382	-----ThrAlaAlaThrAsn
QY	1251	GCCCAATTCAAGGCTCAGGAAGAATCTATGGCAAACTC-----	1301	-----AAGGAGGAGAAC
Db	383	GlyTyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGluLysSer	402	
QY	1302	TTCTTTGGTCCCAAGAGGAGAGTCAAGCTGGAGCCACATACGTGTCCGAGCATCAGCA	1361	
Db	403	ThrAspGlySerLysAspValValGlu-----	418	-----IleAlaValProGluAsnLeu
QY	1362	GCTGGCGGGCTATTGGCAAAAGTGGAAAAACGGTGAACGAGTTGCGAGAATTTGACGGCA	1421	
Db	419	ValGlyAlaIleLeuGlyLysGlyGlyLysThrLeuValGluTyrGlnGluLeuThrGly	438	
QY	1422	GCTGAGGTGGTAGTACCAAGA-----	1475	-----GACCAGACCCCTGATGAGACAGCAGCATCATC
Db	439	AlaArgIleGlnIleSerLysLysGlyGluPheValProGlyThrArgAsnArgLys---	457	
QY	1476	GTCAAAATCATCGACATTTCTATGCCAGTCAGATGGCTCAA	1517	
Db	458	ValThrIleThrGlyThrProAlaAlaThrGlnAlaAlaGln	471	
RESULT 11				
NOA1_HUMAN				
AC	AC	NOA1_HUMAN	STANDARD;	PRT; 510 AA.
DT	DT	P51513;		
DT	DT	01-OCT-1996 (Rel. 34, Created)		
DT	DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	DE	RNA-binding protein Nova-1 (Neuro-oncological ventral antigen 1)		
DE	DE	(Onconeural ventral antigen-1) (Paraneoplastic Ri antigen) (Ventral		
DE	DE	neuron-specific protein 1).		
GN	GN	NOA1		
OS	OS	Homo sapiens (Human)		
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	OX	NCBI_TaxID=9606;		
RN	RN	[1]		
RP	RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).		
RC	RC	TISSUE=Cerebellum, and Hippocampus;		
RX	RX	MEDLINE=34000830; PubMed=8398153;		
RA	RA	Buckanovich R.J., Posner J.B., Darnell R.B.;		
RT	RT	"Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding		
RT	RT	protein and is specifically expressed in the developing motor		
RT	RT	system.";		
FL	FL	Neuron 11:657-672(1993).		
RP	RP	[2]		
RP	RP	SEQUENCE OF 1-34 FROM N.A.		
RC	RC	TISSUE=Fetal brain;		
RA	RA	Dmitrenko V.V., Garifulin O.M., Shostak K.A., Smikodub A.I.		

Pred. No.:	3.09e-07	Length:	707
Score:	234.50	Matches:	144
Percent Similarity:	32.93%	Conservative:	74
Best Local Similarity:	21.75%	Mismatches:	220
Query Match:	6.64%	Indels:	224
DB:	1	Gaps:	29

US-09-270-437D-7 (1-1946) x FUB2_HUMAN (1-707)

QY	275	AGCCATCATGAAGCTGAATGGCCACCACTGGGAACCACTGCCCTGAAGGTCTCTCATACAT	334
DB	2	SerAspTyrSerThrGlyGlyProProGlyProProGlyProProProAlaGly-----	18
QY	335	CCCCGATGACGAGATACACACAGGACCTGAGATGGCGCGCGGGGCTTGGCTCTCG	394
DB	19	-----GlyGlyGlyGlyAlaGlyGlyAla	26
QY	395	GGGTCCAGCCCGCCA-----GGGCTCACCTGT	421
DB	27	GlyGlyGlyProProProGlyProProGlyAlaGlyAspArgGlyGlyGlyProCys	46
QY	422	GGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGA-----CATCCCCCTTGGCTCTCT	475
DB	47	Gly---GlyGlyProGlyGlyGlySerAlaGlyGlyProSerGlnProProGlyGlyGly	65
QY	476	GGTSCCC-----ACCAGTATGGGG	496
DB	66	Gly---ProGlyIleArgIysAspAlaPheAlaAspAlaValGlnArgAlaGlnIleAl	85
QY	497	TGCCATTATTGGCAAGGAGGGGCCCATCCATCCGCAACATCACAAA-----	542
DB	85	aAlaIysIleGlyGlyAspAlaAlaThrThrGlyAsnAsnSerThrProAspPheGlyPh	105
QY	543	-----CAGACCCAGTCCAGATAGAGCTGCATAGGACGAGAACGACGAGTGCAGCTGA	595
DB	105	eGlyGlyGlnIysArgGlnLeuGluAspGlyAspGlnProGlnSerIysIysLeuAlaSe	125
QY	596	AAAAGCCATCATGTGCACTCCACCCCTCAGGGGCTGCTCCCTCGCTTGTAAAGATGATCTT	655
DB	125	rGlnGlyAspSerIleSerSerGln-----Le	134
QY	656	GGAGATTATGCATTAAGAGGCTAAGACACCAAAACGGCTGCAC-----GAGGTTCCCTGAA	712
DB	134	uGlyProIleHisProProProArgThrSerMetThrGluGluTyrArgValPro-----	152
QY	713	GATCCTGSCCCATAACTTTGTAGGGCGTCTCATTGGCAAGCAAGCAGCAACCTGAA	772
DB	153	-----AspGlyMetValGlyLeuIleIleGlyArgGlyGlyGluGlnIleAs	168
QY	773	GAAGGTAGACGATACCGAGACAAAAATCACCATCTCC-----TCGTTGCA	820
DB	168	nLysIleGlnAspSerGlyCysLysValGlnIleSerProAspSerGlyGlyLeu--	187
QY	821	AGACTTACCTTTACACCTGACAGGACCATCACTGTGAAGGGGCCATCGAGAATTG	880
DB	188	-----ProGluArgSerValSerLeuThrGlyAlaProGluSerVa	201
QY	881	TTGACGGGCCGACGAGGAATAATGAAGAAAGTTTCGGAGGCCCTATGAGAATGATGTGC	940
DB	201	lGlnLysAlaLysMetLeu-----AspAspIleVa	212
QY	941	TGCCATGAGCTCTACCTGATCCCTGGCTGAACTGGCTGCTGTAGTCTTTTCCGAGC	1000
DB	212	IserArgGlyArgGlyProProGly-----	221
QY	1001	TTCATCCAGCGCAGTCCCGCGCCTCCACAGACGCGTTACTGGGCTGTCTCCCTATAGCTC	1060
DB	222	-----GlnPheHisAs	225
QY	1061	CTTTATGAGGCTCCGAGCAGGAGATGGTCAG---GTGTTTATPCCCGCCGAGGCAGT	1117
DB	225	pAsnAlaAsnGlyGlyGlnAsnGlyThrValGlnGluIleMetIleProAlaGlyLysAl	245

RESULT 8

RA Yang Y.Y., Yin G.L., Darnell R.B.:
 RT "The neuronal RNA-binding protein Nova-2 is implicated as the
 RT autoantigen targeted in POMA patients with dementia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13254-13259(1998).
 [3]
 RP SEQUENCE OF 29-492 FROM N.A.
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,
 RA Burkhardt-Schultz K.J., Gordon L., Dias J., Ramirez M., Stillwagen S.,
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,
 RA Danganan L., Erlar A., Christensen M., Georgescu A., Avila J., Liu S.,
 RA Attix C., Andreise T., Frankheim M., Amico-Keller G., Coesfield J.,
 RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
 RA Arellano A., Brower A., Sanders C., Ow D., Nolan M., Trong S.,
 RA Kobayashi A., Olsen A.S., Carrano A.V.:
 RA "Sequence analysis of a 1.9 Mb region in 19q13.2 between APOE and
 RT D19S412.";
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 405-480.
 RX MEDLINE=99148126; PubMed=10368286;
 RA Lewis H.A., Chen H., Edo C., Buckanovich R.J., Yang Y.Y.-L.,
 RA Musunuru K., Zhong R., Darnell R.B., Burley S.K.:
 RT "Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding
 RT domains.";
 RL Structure 7:191-203 (1999).
 [5]
 RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 405-491 IN COMPLEX WITH RNA.
 RX MEDLINE=20139984; PubMed=10676814;
 RA Lewis H.A., Musunuru K., Jensen K.B., Edo C., Chen H., Darnell R.B.,
 RA Burley S.K.:
 RT "Sequence-specific RNA binding by a Nova KH domain: implications for
 RT paraneoplastic disease and the fragile X syndrome.";
 RL Cell 100:323-332(2000).
 CC -!- FUNCTION: May regulate RNA splicing or metabolism in a specific
 CC subset of developing neurons (By similarity). Binds single strand
 CC RNA.
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- TISSUE SPECIFICITY: Brain. Expression restricted to astrocytes.
 CC -!- DOMAIN: The third KH domain (KH3) recognizes specifically 5'-
 CC UCAY-3'.
 CC -!- DISEASE: Autoantigen in the paraneoplastic opsoclonus myoclonus
 CC ataxia (POMA), a paraneoplastic neurological syndrome/disorder
 CC (PNS/P) associated with breast cancer, fallopian cancer, and SCLCa
 CC and characterized primarily by loss of inhibitory control of motor
 CC neurons in the spinal cord and brainstem. Recognized by the IgG
 CC autoantibody ANNA-2 (also called anti-R1).
 CC -!- SIMILARITY: Contains 3 KH domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U70477; AAB88661.1; ALT_INIT.
 DR EMBL; AF083898; AAC72355.1; -.
 DR EMBL; AC006540; AAD13116.1; -.
 DR PDB; 1DTJ; 18-FEB-00.
 DR PDB; 1EC6; 07-APR-00.
 DR Genew; HGNC:7887; NOVA2.
 DR MIM; 601991; -.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 3.
 DR SMART; SM00322; KH; 3.
 DR PROSITE; PS00084; KH_TYPE_1; 3.
 KW Antigen; Nuclear protein; RNA-binding; Repeat; 3D-structure.
 FT DOMAIN 10 26 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
 FT FT
 FT DOMAIN 32 99 KH 1.
 FT FT
 FT DOMAIN 130 196 KH 2.
 FT FT

FT	DOMAIN	232	405	ALA-RICH.
FT	DOMAIN	332	393	GLY-RICH.
FT	DOMAIN	406	473	KH 3.
FT	CONFLICT	247		A -> R (IN REF. 1).
FT	CONFLICT	265	267	PAA -> TAT (IN REF. 1).
SQ	SEQUENCE	492 AA;	49008 MW;	41B63EAF6899256B CRC64;

Alignment Scores:				
Pred. No.:	2.12e-07	Length:	492	
Score:	236.50	Matches:	118	
Percent Similarity:	37.89%	Conservative:	62	
Best Local Similarity:	24.84%	Mismatches:	185	
Query Match:	6.70%	Indels:	111	
DB:	1	Gaps:	15	

US-09-270-437D-7 (1-1946) x NOA2_HUMAN (1-492)

QY	390	TCTGGGGTCCAGCCCGCCGCGGCTCACCT	-----GTGGCAGCGGGGGCCCGCCAG	443
DB	8	SerArgLysArgProLeuGluThrProGluValValCysThrLysArgSerAsnThr	27	
QY	444	CAGCAGCAAGTGGACATCCCTTCGGCTCTCTGGTGGCCACCAGTANGTGGTGGCCATT	503	
DB	28	GlyGluGluGlyGluTyrPheLeuLysValLeuLeuProSerTyrAlaAlaGlySerIle	47	
QY	504	ATTGGCAGAGGAGGGCCACCATCCGCAACATCCAAACACAGCCACGATCCCAAGATAGAC	563	
DB	48	IleGlyLysGlyGlyGlnThrIleValGlnLeuGlnLysGluThrGlyAlaThrIleLys	67	
QY	564	GTGCATAGCAAGAGGAAC-----GCAGGTGCAGCTGAAAGAACCATCATGTGTGCATCC	617	
DB	68	LeuSerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuValGlnGly	87	
QY	618	ACCCCTGAGGGCTGCTCTCCGCTTGTGAAGATGATCTTGGAGATATATGATATAAGAGGCT	677	
DB	88	ThrAlaGluAlaLeuAsnAlaValHisSerPheIleAlaGluLysVal---ArgGluIle	106	
QY	678	AAGGACACCAAAACGGCTGACGAGGTT	704	
DB	107	ProGlnAlaMetThrLysProGluValValAsnIleLeuGlnProGlnThrThrMetAsn	126	
QY	705	-----CCCTCGAAGATCTCGGCCATAATACTTTGTAGGGCGTCTCAT	749	
DB	127	ProAspArgAlaLysGlnAlaLysLeuIleValProAsnSerThrAlaGlyLeuIleIle	146	
QY	750	GCACAGGAGGACGGAACCTGAGAGGTAGGACAGATACCGACACAAATACCATC	809	
DB	147	GlyLysGlyGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpValGlnLeu	166	
QY	810	TCC-----TCGTTGCAAGACCTTACCTTTACACCTCAGAGGACC	851	
DB	167	SerGlnLysProGluGlyIleAsnLeuGln-----GluArgVal	179	
QY	852	AFCACGTCTGAAGGGGCCCATCGAGAAATTGTCAGGGCGGACGAGAAATATGAAGAA	911	
DB	180	ValThrValSerGlyGluProGluGlnValHisLysAlaValSerAlaIleValGlnLys	199	
QY	912	GTTCCGGAG-----GCCTATCAGAAAT	932	
DB	200	ValGlnGluAspProGlnSerSerCysLeuAsnIleSerTyrAlaAsnValAlaGly	219	
QY	933	-----GATGTG-----	938	
DB	220	ProValAlaAsnSerAsnProThrGlySerProTyrAlaSerProAlaAspValLeuPro	239	
QY	939	-----GCTCCCATGAGCTCTACCTGATCCCTGGCTGAGCTGAGCTGGCT	980	
DB	240	AlaAlaAlaAlaAlaSerAlaAlaAlaAlaSerGlyLeuLeuGlyProAlaGlyLeuAla	259	
QY	981	GCTGTAGGTCTTTCCAGCTTC-----ATCCAGCGC	1012	
DB	260	GlyValGlyAlaPheProAlaAlaLeuProAlaPheSerGlyThrAspLeuLeuAlaIle	279	

```
QY 332 CATCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGGGCGCGAGGGGGCTTTGGGCTC 391
Db 171 ---ProArg-----
QY 392 TCGGGGTGAGCCCGCCAGGGCTCACCTGTGGC-----AGCGGGGGCCCGCAGC 439
Db 177 ---ArgArgProProAlaLeuThrGlyGlyAlaLeuProSerAlaLeuProPro 195
QY 440 CAAGCAGCAGCAAGTGGACATCCCTTCGG-----CTCCTGGT 478
Db 196 GlnLeuGlyProMetHisProProPro-ArgSerThrValThrGluGluTyrArgVa 215
QY 479 GCCCACCAGTATGTGGTGGCCATTATTGGCAAGGAGGGGCCACCATCCCAACATCAC 538
Db 215 lProAspGlyMetValGlyLeuIleIleGlyArgGlyGlyGlnIleAsnIleGly 235
QY 539 AAAACACACCCAGTCCCAAGATAGACGTGCATAGGAGGAGAACCGCAGCTGCAAAA 598
Db 235 nGlnAspSerGlyCysIysValGlnIle---SerProAspSerGlyGlyLeuProGluAr 254
QY 599 AGCCATCAGTGTGACCTCCACCCCTGAGGGCTGCTCCTCCGCTTGTAAAGATGATCTTGA 658
Db 254 gSerValSerLeuThrGlySerProGluAlaValGlnLysAlaLysLeuMetLeuAspAs 274
QY 659 GATTATG-----CATAAAGAGGCTAAG---GA 682
Db 274 pIleValSerArgGlyArgGlyGlyProProGlyGlnPheHisAspTyrAlaAsnGlyGl 294
QY 683 CACCAAAACGGCTGACGAGGTTCCTCGAAGATCTCGCCCATATTAATTTGTAGGGCG 742
Db 294 nAsnGlyThrValGlnGlu-----IleMetIleProAlaGlyLysAlaGlyLe 310
QY 743 TCTCATTTGCAAGGAAGCGAGCACTGCAAGAGGTAGAGCAAGATACCGAGACAAAAT 802
Db 310 uValIleGlyIysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValIysMe 330
QY 803 CACCATC-----TCCTCGTTGCAAGACCTTACCCTTTACAACCTCAGAGGACCATCAC 856
Db 330 tIlePheIleGlnAspGlySerGlnAsnThrAsnValAspLysProLeuArgIleIleGl 350
QY 857 TGTGAAGGGGCCATCCAGATTTGTCAGGGCCGAGCAGCAAGAAATATGAAGATTCG 916
Db 350 yAspProTyrLysValGlnGlnAlaCysGluMetValMetAspIleLeuArgGluArgAs 370
QY 917 GGAGGCTATCAGAAATGATGTGGCTGCCATCAGCTCTCACCTGATCCTCGCTGCAACCT 976
Db 370 pGlnGlyGlyPheGlyAspArgAsnGluTyrGlySerArgIleGlyGlyIleAsp-- 389
QY 977 GGCTGTGTAGTCTTTCCAGCTTCATCCAGGGCAGTCCCGCCGCTCCCGCAGCAGCT 1036
Db 390 -----ValPro----- 391
QY 1037 TACTGGGGCTGCTCCCTATAGCTCTTTATGCAGGCTCCCGAGCAGGAGATGGTGAGGT 1096
Db 391 ----- 391
QY 1097 GTTTATCCCGCCGAGCAGTGGCGGCATCATCGGCAAGAGGGCGAGCAGCATCAACA 1156
Db 392 ---ValProArgHisSerValGlyValValIleGlyArgSerGlyGluMetIleLysLy 410
QY 1157 GCTCTCCCGTTGGCCAGCCCTCCATC----- 1184
Db 410 sIleGlnAsnAspAlaGlyValArgIleGlnPheLysGlnAspAspGlyThrGlyProGl 430
QY 1185 -AAGATTGCA-----CCACCCGAAACACCTGACTCCCAAAATTCGTATGGTT-- 1229
Db 430 uLysIleAlaHisIleMetGlyProProGluArgCysGluHisAlaAlaArgIleIleAs 450
QY 1230 -----ATCATCATGTCGACCGCCAGAGGCCCAATT----- 1258
Db 450 nAspLeuLeuGlnSerLeuArgSerGlyProPro--GlyProProGlyHisGlyMetPro 469
```

```
QY 1259 -----CAAGGCTCAGGAAGAATCTATGGCAAACTCAAGAGGAGAACTT 1303
Db 470 ProGlyGlyArgGlyArgGlyGlyGlnGlyIleTrpGlyProProGlyGlyGlyMet 489
QY 1304 CTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCACATACGTGTGCCAGCATCAGCAGC 1363
Db 490 ThrPhe-----Ser-IleProThrHisLysCy 498
QY 1364 TGGCCGGGTTCATTGGCAAGGTGAAAAACGGTGAACAGATTGCAGAAATTTGACGGCAGC 1423
Db 498 sGlyLeuValIleGlyArgGlyGlyGluAsnValLysAlaIleAsnGlnArgGlyAl 518
QY 1424 TGAGTGTGTAGTACCAAGAGACACCCCT-----GATGAGAACACGACGATCATCGT 1477
Db 518 aPheValGluIleSerArgGlnLeuProProAsnGlyAspProAsnPheLysLeuPheI 538
QY 1478 GAAATCATCGGACATTCTATGCCAGTCAGATGGCTCAACGGGAAGATCCGAGACATCCT 1537
Db 538 e-----IleArgGlySerPr 543
QY 1538 GGCCCGAGTTAAGCAGCAGCATCAGAGGAGGACAGAGTAACCGAGGCCCGCAGCAGGAA 1597
Db 543 oGlnGlnIleGluHisAlaLysGln-----ProIleGluGluLy 556
QY 1598 GTGACGAGCCCTCCCTGCTCCCTTNGAGTCCAGGACAAACACGGGCGAGAAATCGAGATG 1657
Db 556 sIleGluGlyProLeuCysPro----- 563
QY 1658 TGCTCTCCCGCGCAGGCTGAGATGAGTGGGAATCCGGGACACNTGGCGGCTGTAG 1717
Db 564 -----ValGlyProGlyProGl 569
QY 1718 ATCAGGTTTCCACATTGATTGAGAAAGATGTTCCAGTGAAGAACCTGATCTTCAGCC 1777
Db 569 yProGlyGlyProPro----- 574
QY 1778 CCAACACCCACCAACCAATGGCCCAACTGNTGCCCTCGGGGTGTCAGAAATTTAGC 1837
Db 575 -----GlyProAlaGly-----Pr 579
QY 1838 GCAAGGCACCTTTAAACGTGATTTTAAAGAGCT-----CT 1876
Db 579 oMetGlyProPheAsnProGlyProPheAsnGlnGlyProProGlyGlyProProHi 599
QY 1877 CCAAGGCCCAACCAAGAGGTGG-----ATCACACCTCAGTGGGAA 1916
Db 599 sGlnTyrProProGlnGlyTrpGlyAsnThrTyrProGlnTrpGln 614

RESULT 6
NOA2 HUMAN STANDARD; PRT; 492 AA.
AC Q9UNW9; O43267; Q9UEA1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE RNA-binding protein Nova-2 (Neuro-oncological ventral antigen 2)
DE (Astrocytic NOVA1-like RNA-binding protein).
DE NOVA2 OR ANOVA OR NOVA3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20137319; PubMed=10735272;
RA Ueki K., Ramaswamy S., Billings S.J., Mohrenweiser H.W., Louis D.N.;
RT "ANOVA, a putative astrocytic RNA binding protein gene that maps to
RL chromosome 19q13.3.";
RL Neurogenetics 1:31-36 (1997).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=99007301; PubMed=9789075;
```


QY 1000 CTTTCATCCAGCGAGTCCCGCGCTCCACGAGCGTTACTGGGCTCTCCCTATAGCT 1059
 Db 320 -----SerArgValGlyGly----- 325
 QY 1060 CTTTATGCGAGTCCCGAGGAGATGGTCAGGTGTTTATCCCGCCAGGCGAGTGG 1119
 Db 326 -----IleAspValProValProArgHisSerValG 336
 QY 1120 GCGCATCATCGCAAGAGGGCGACACATCAACAGCTCTCCCGTTTCCGACGCGCT 1179
 Db 336 lyValValIleGlyArgSerGlyGluMetIleLysLysIleGlnAsnAspAlaGlyVala 356
 QY 1180 CCATC-----AGATTGCA-----C 1194
 Db 356 rgIleGlnPheLysGlnAspAspGlyThrGlyProGluLysIleAlaHisIleMetGlyP 376
 QY 1195 CACCGGAACACCTCACTCCAAAGTTCGTATGGT-----ATCATCA 1236
 Db 376 roProAspArgCysGluHisAlaAlaArgIleIleAsnAspLeuGlnSerLeuArgS 396
 QY 1237 CTGGACCCCGAGGCGCAATT-----CAAGG 1263
 Db 396 erGlyProPro--GlyProGlyAlaProGlyMetProGlyGlyArgGlyArgG 415
 QY 1264 CTCAGGAGATCATATGGCAACTCAAGGAGAGACTTCTTTGGTCCAGGAGGAGAG 1323
 Db 415 yArgGlyGlnGlyAsnTrpGlyProGlyGlyGluMetThrPhe----- 430
 QY 1324 TGAAGCTGGAGACACATATCGTGTGCCAGCATCAGCAGCTGGCGGGTCATTGGCAAG 1383
 Db 431 -----Ser-IleProThrHisLysCysGlyLeuValIleGlyArgG 444
 QY 1384 GTGGAATAACCGTGAACGATTCAGAAATTCACGCGAGCTAGGTGGTAGTACCAAGAG 1443
 Db 444 lyGlyGluAsnValLysAlaIleAsnGlnGlnThrGlyAlaPheValGluIleSerArgG 464
 QY 1444 ACCAGACCCCT-----GATGAGACGACCATCGTCTATCGT-----AAAA 1482
 Db 464 lnLeuProProAsnGlyAspProAsnPheLysLeuPheValIleArgGlySerProGlnG 484
 QY 1483 TCATCGGACATTTCTATGCGAGTTCAGATGGCTCAACGAGGATCCGACAGATCTTGGCC 1542
 Db 484 lnIleAspHis-----AlaLysGlnLeuIleGluLysIleGlu----- 497
 QY 1543 AGGTTAAGCAGCATCATCAAGAGGACAGAGTAACGAGGCCCGGACGAGGAGAGTAGAC 1602
 Db 497 ----- 497
 QY 1603 CAGCCCTCCCTGTCCTTNGAGTCCAGGACACACGCGGACAAATCGAGAGTGTGCTC 1662
 Db 498 --GlyProLeuCysPro----- 502
 QY 1663 TCCCCGCGAGCGCTGAGATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAG 1722
 Db 503 -----ValGlyProGly----- 506
 QY 1723 GTTTGCCCATTTGANTGAGAAAGATGTTTCAGTGAGAACCTGTACTNTCAGCCCCAAA 1782
 Db 506 ----- 506
 QY 1783 CACCCACCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTGAGAAATTTAGGCAAG 1842
 Db 507 -----ProGlyGlyProGlyProAlaGlyProMetG 517
 QY 1843 GCACCTTTTAAACGTGATTGTTTAAAGAGCT----- 1874
 Db 517 lyProPheHisProGlyProPheAsnGlnGlyProGlyAlaProProHisAlaGlyG 537
 QY 1875 -----CTCCAGGCCCAACCAAGAGGGTGG-----ATCACACCTCATGTGGAA 1916
 Db 537 lyProProHisGlnTyrProProGlnGlyTrpGlyAsnThrTyrProGlnTrpGln 556

RESULT 5

ZBP2_CHICK
 ID ZBP2_CHICK STANDARD; PRT; 769 AA.
 AC Q8UVD9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zipcode-binding protein 2.
 GN ZBP2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBT_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 104-119; 242-267; 679-685; 705-711 AND
 RP 743-766, FUNCTION, DEVELOPMENTAL STAGE, AND SUBCELLULAR LOCATION.
 RX MEDLINE=21640447; PubMed=11781334;
 RA W., Pan F., Zhang H., Bassell G.J., Singer R.H.;
 RT "A predominantly nuclear protein affecting cytoplasmic localization of
 RT beta-actin mRNA in fibroblasts and neurons.";
 RL J. Cell Biol. 156:41-52(2002).
 CC -!- FUNCTION: Binds to a 54-nucleotide localization signal (the
 CC zipcode) found in the 3' untranslated region of beta-actin mRNA
 CC and may play a role in its localization.
 CC -!- SUBCELLULAR LOCATION: Predominantly nuclear. Also detected in the
 CC cytoplasm.
 CC -!- DEVELOPMENTAL STAGE: The highest expression is found in 6-d
 CC embryos, is reduced to 30% before hatching and remains stable
 CC thereafter.
 CC -!- SIMILARITY: Contains 4 KH domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF461020; AAL66365.1; -;
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0003730; F:mRNA 3' UTR binding; IDA.
 DR GO; GO:0008298; P:mRNA localization, intracellular; NAS.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 4.
 DR SMART; SM00322; KH; 4.
 DR PROSITE; PS50084; KH TYPE 1; 4.
 DR Transport; mRNA transport; Nuclear protein; RNA-binding; Repeat.
 KW DOMAIN 208 272 KH 1.
 FT DOMAIN 296 362 KH 2.
 FT DOMAIN 385 449 KH 3.
 FT DOMAIN 486 553 KH 4.
 SQ SEQUENCE 769 AA; 80644 MW; 875A6C83529969EB CRC64;
 Alignment Scores:
 Pred. No.: 2,05e-07 Length: 769
 Score: 237.50 Matches: 135
 Percent Similarity: 32.29% Conservative: 71
 Best Local Similarity: 21.16% Mismatches: 207
 Query Match: 6.73% Indels: 225
 DB: 1 Gaps: 25
 US-09-270-437D-7 (1-1946) x ZBP2_CHICK (1-769)
 QY 212 GAACACCGAGTGAGACGCGAGTGGTGAATGTACCTATTCCACCCGGGAGCAGCAG 271
 Db 135 GluSerLysLysLeuAlaGlnGlyAspCysGluTyrGlyProArgGlyGlyAlaPro 154
 QY 272 GCAAGCCATCATGAAGCTGAATGGCCACCGAGTTGGAGAACCATGCGCTCGAAGTCTCTTA 331
 Db 155 AlaAlaProGluArgSerGlyProValGlyAspProGlyPro----- 170

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Far upstream element binding protein 2 (FUSE binding protein 2) (KH
 DE type splicing element binding protein) (KSRP) (MAP2 RNA trans-acting
 DE protein 1) (MARTAL1).
 GN FUBP2 OR KHSRP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., SEQUENCE OF 73-87; 89-108 AND 475-486, SUBCELLULAR
 RP LOCATION, AND FUNCTION.
 RC TISSUE=Brain;
 RX MEDLINE=22246918; PubMed=12358751;
 RA Rehbein M., Wege K., Buck F., Schweizer M., Richter D., Kindler S.,
 RT "Molecular characterization of MARTAL, a protein interacting with the
 RT dendritic targeting element of MAP2 mRNAs".
 RL J. Neurochem. 82:1039-1046(2002).
 CC -!- FUNCTION: Part of a ternary complex that binds to the downstream
 CC control sequence (DCS) of the pre-mRNA. Mediates exon inclusion in
 CC transcripts that are subject to tissue-specific alternative
 CC splicing. May interact with single-stranded DNA from the far-
 CC upstream element (FUSE). May activate gene expression (by
 CC similarity). Binds to the dendritic targeting element and may play
 CC a role in mRNA trafficking.
 CC -!- SUBUNIT: Part of a ternary complex containing FUBP2, PTBP1, PTBP2
 CC and HNRPH1 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear. A small proportion is also found in
 CC the cytoplasm of neuronal cell bodies and dendrites.
 CC -!- SIMILARITY: Contains 4 KH domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: AF308818; AAG59811.1; -.
 DR InterPro: IPR004087; KH dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam: PF00013; KH; 4.
 DR SMART: SM00322; KH; 4.
 DR PROSITE: PS50084; KH TYPE 1; 4.
 KW Transport; mRNA transport; mRNA processing; mRNA splicing;
 KW Transcription regulation; Trans-acting factor; Nuclear protein;
 KW DNA-binding; RNA-binding; Repeat.
 FT DOMAIN 145 209
 FT DOMAIN 234 300 KH 1.
 FT DOMAIN 323 387 KH 2.
 FT DOMAIN 425 492 KH 3.
 FT DOMAIN 572 685 KH 4.
 FT REPEAT 572 583 4 X 12 AA IMPERFECT REPEATS.
 FT REPEAT 618 629 1.
 FT REPEAT 644 655 2.
 FT REPEAT 674 685 3.
 FT REPEAT 69 68 GLY/PRO-RICH.
 FT DOMAIN 69 498 GLY-RICH.
 FT DOMAIN 499 613 ALA/GLY/PRO-RICH.
 SQ SEQUENCE 721 AA; 74226 MW; 482C7A765C60E84A CRC64;

Alignment Scores:

Pred. No.:	7.36e-08	Length:	721
Score:	244.50	Matches:	151
Percent Similarity:	31.26%	Conservative:	75
Best Local Similarity:	20.89%	Mismatches:	221
Query Match:	6.92%	Indels:	276
DB:	1	Gaps:	29

US-09-270-437D-7 (1-1946) x FUB2_RAT (1-721)

QY	52	GGCCTCAGCCACCCAGAGCGCGGGTGGAGGGCGAGTGTCTCAGCTTCCGGGTAGGA	111
Db	12	GlyProProProProAlaGlyGlyGlyGlyGlyAlaAa	25
QY	112	GCCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGAGTACTGGACGCC	171
Db	26	AlaGly-GlyGlyProPro	41
QY	172	TGCTGGCTCAGTATGGTACAGTAGAAGTCTGTAGCAAGTGAACACCGAGAGTGAGCGG	231
Db	41	YGLYGLYGLY	46
QY	232	CAGTGGTGAATGTCACTATTCCAAACCGGGAGCAGACAGCCATCATGAAGTGA	291
Db	46	YGLYGLY	61
QY	292	ATGGCCACCATGGAGAACCATGCCCTGAAGGTCTCTCATATCCCGAT	341
Db	61	nProProGlyGlyGlyGlyPro-GlyIleArgLysAspAlaPheAlaAspAlaValGlnA	81
QY	342	-----GAGCAGATAGCACAG-----GGACCTGAG-----AATGGCGGCC	375
Db	81	rgAlaAa-rgGlnIleAlaAlaLysIleGlyGlyAspAlaAlaThrValAsnAsnAsnT	101
QY	376	GAGGGGCTTTGGCTCTCGGGTCAGCCCGCCAG-----	410
Db	101	hrProAspPheGlyGlyGlyGlnLysArgGlnLeuGluAspGlyAspGlnProAsps	121
QY	411	-----GCTCAGCTGTGGCAGCGGGGGCCAG	438
Db	121	erLysLysLeuAlaSerGlnGlyAspSerIleGlySerGlnLeuGlyProIleHisProp	141
QY	439	CCAAGCAGCAGCAGTGCATCCCTCGGCTCTGGTCTCCGCCACCCAGTATGGGTG	498
Db	141	roProAa-rgThrSerMetThrGluGluTyrArg-----ValProAspGlyMetValGlyL	159
QY	499	CCATTATTGCAAGAGGGGCCACCATCCGCAACATCACAAAACAGCAGCCAGTCCAAGA	558
Db	159	eulleIleGlyArgGlyGlyGluGlnIleAsnLysIleGlnGlnAspSerGlyCysLysv	179
QY	559	TAGACGTGCATAGGAGAGAACCGCAGTGCAGTGAAGAACCCATCAGTGTGCACTCCA	618
Db	179	alGlnIle-----SerProAspSerGlyGlyLeuProGluArgSerValSerLeuThrGlyA	198
QY	619	CCCTGAGGCTGCTCTCCGCTGTGAAGATGATCTTGGAGATTATG-----	665
Db	198	laProGluSerValGlnLysAlaLysMetMetMetMetMetMetMetMetMetMetMet	218
QY	666	-----CATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGGTTC	705
Db	218	lyGlyProProGlyGlnPheHisAspAsnAlaAsnGlyGlyGlnAsnGlyThrValGlnG	238
QY	706	CCCTGAGATCTCGGCCCATATACTTTGTAGGGCTCTCATTTGGCAAGAGGACGGA	765
Db	238	luIleMetIleProAlaGlyLys-----AlaGlyLeuValIleGlyGlyGlyGluT	256
QY	766	ACCTGAAGAGGTAGAGCAGATACCGAGACAAAATCACCATC-----TCCTCGTTC	819
Db	256	hrIleLysGlnLeuGlnGluArgAlaGlyValLysMetIleLeuIleGlnAspGlySerG	276
QY	820	AAGACCTTACCTTTACACCTCGAGAGGACCATCTGTGAAGGGGGCCCATCGAGAATT	879
Db	276	lnAsnThrAsnValAspLysProLeuArgIleIleGlyAspProTyrLysValGlnGlnA	296
QY	880	GTTGAGGGCCGAGCAGGAATATGAAGAAATTCGGGAGGGCTATGAGAAATGATGG	939
Db	296	laCysGluMetValMetAspIleLeuArg-----GluArgAspGlnG	310
QY	940	CTGCCATGAGCTCTCACCTGATCCCTGGCTGAACCTGGCTGTGTAGGTCTTTTCCAG	999
Db	310	lyGlyPheGlyAspArgAsnGluTyrGly-----	319


```
QY 1813 CCCTCGGGTGTGAGAAATNTAGCGGACGACCTTTAAACGTGGATTGTTTAAAGAG 1872
Db 560 nThrAsnGlyGlnGlyAspGlnGlnAla----- 569
QY 1873 CTCTCCAGGCCCCACCAAGAGGGTGGATCACACC-----TC 1908
Db 570 ---ProAlaProAlaGlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysLys 588
QY 1909 AGTCGGAGAAATAATAAATTCCTTCAGGT 1939
Db 588 sMetGlyGlnAlaValProAlaProAlaGly 598

RESULT 3
FUBI_HUMAN
ID FUBI_HUMAN STANDARD; PRT; 643 AA.
AC Q96AE4; Q12828;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Far upstream element binding protein 1 (FUSE binding protein 1) (FBP)
DE (DNA helicase V) (HDH V).
GN FUBP1.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), SEQUENCE OF 191-193; 203-206;
RP 272-279; 284-290; 300-314; 321-328; 394-397; 409-411; 430-438 AND
RP 440-443, AND FUNCTION.
RC TISSUE=Leukemia;
RX MEDLINE=94170991; PubMed=9125259;
RA Duncan R., Bazar L., Michelotti G., Tomonaga T., Krutzsch H.,
RA Avigan M., Levens D.;
RT "A sequence-specific, single-strand binding protein activates the far
RT upstream element of c-myc and defines a new DNA-binding motif.";
RL Genes Dev. 8:465-480(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 45-64; 133-145; 271-283; 308-321; 379-386; 414-424 AND
RP 430-439.
RX MEDLINE=21127960; PubMed=11222755;
RA Vindigni A., Ocham A., Triolo G., Falaschi A.;
RT "Identification of human DNA helicase V with the far upstream
RT element-binding protein.";
RL Nucleic Acids Res. 29:1061-1067(2001).
RN [4]
RP PARTIAL SEQUENCE, AND MASS SPECTROSCOPY.
RX MEDLINE=22166132; PubMed=12176931;
```

```
RA Rappsilber J., Ryder U., Lamond A.I., Mann M.;
RT "Large-scale proteomic analysis of the human spliceosome.";
RL Genome Res. 12:1231-1245(2002).
RN [5]
RP INTERACTION WITH SIAHBP1.
RX MEDLINE=20337922; PubMed=10882074;
RA Liu J., He L., Collins I., Ge H., Libutti D., Li J., Egly J.-M.,
RA Levens D.;
RT "The FBP interacting repressor targets TFIID to inhibit activated
RT transcription.";
RL Mol. Cell 5:331-341(2000).
RN [6]
RP INTERACTION WITH JTV1, UBIQUITINATION, AND PROTEASOME-MEDIATED
RP DEGRADATION.
RX MEDLINE=22716800; PubMed=12819782;
RA Kim M.-J., Park B.-J., Kang Y.-S., Kim H.J., Park J.-H., Kang J.W.,
RA Lee S.W., Han J.M., Lee H.-W., Kim S.;
RT "Downregulation of FUSE-binding protein and c-myc by tRNA synthetase
RT cofactor p38 is required for lung cell differentiation.";
RL Nat. Genet. 34:330-336(2003).
RN [7]
RP STRUCTURE BY NMR OF 278-447 IN COMPLEX WITH SINGLE STRANDED FUSE DNA.
RX MEDLINE=21864720; PubMed=11875576;
RA Braddock D.T., Louis J.M., Baber J.L., Levens D., Clore G.M.;
RT "Structure and dynamics of KH domains from FBP bound to
RT single-stranded DNA.";
RL Nature 415:1051-1056(2002).
CC -!- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC far-upstream element (FUSE) upstream of the MYC promoter. May act
CC both as activator and repressor of transcription.
CC -!- SUBUNIT: Interacts with SIAHBP1/FIR and JTV1.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC Name=2;
CC IsoId=Q96AE4-1; Sequence=Displayed;
CC IsoId=Q96AE4-2; Sequence=VSP_008321;
CC Note=No experimental confirmation available;
CC -!- PTM: Ubiquitinated. This targets the protein for proteasome-
CC mediated degradation.
CC -!- SIMILARITY: Contains 4 KH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U05040; AAA17976.2; -.
CC EMBL; BC017247; AAH17247.1; -.
CC PDB; 1J4W; 06-MAR-02
CC Genew; HGNC:4004; FUBP1.
CC MIM; 603444; -.
CC InterPro; IPR008160; Collagen.
CC InterPro; IPR004087; KH dom.
CC InterPro; IPR004089; KH_type_1.
CC Pfam; PF00013; KH; 4.
CC SMART; SM00322; KH; 4.
CC PROSITE; PS50084; KH TYPE 1; 4.
CC Transcription regulation; Trans-acting factor; Nuclear protein;
CC DNA-binding; Repeat; Ubl conjugation; Alternative splicing;
CC 3D-structure.
CC 99 163 KH 1.
CC DOMAIN 184 250 KH 2.
CC DOMAIN 274 338 KH 3.
CC DOMAIN 375 442 KH 4.
CC DOMAIN 13 26 GLY-RICH.
CC DOMAIN 348 395 GLY-RICH.
CC DOMAIN 449 559 PRO-RICH.
CC VARSPLIC 642 643 GQ -> CRDPASIELAL (in isoform 2).
```

2).
 /FTId=VSP_008322.
 MISSING (IN REF. 2).
 FT CONFLICT 93 93
 SQ SEQUENCE 651 AA; 68539 MW; F4EACA62FD9FA0D5 CRC64;

Alignment Scores:
 Pred. No.: 2.81e-08 Length: 651
 Score: 251.00 Matches: 157
 Percent Similarity: 35.06% Conservative: 93
 Best Local Similarity: 22.02% Mismatches: 237
 Query Match: 7.11% Indels: 227
 DB: 1 Gaps: 34

US-09-270-437D-7 (1-1946) x FUBI_MOUSE (1-651)

QY	134	TATTCCACC-----	CCAGCTCCGATGGGAAGTACTGACAGCCGCTGCTGGCTCAGTATGG	187	
Db	4	TyrSerThrValProProSerSerGlySerAlaGlyGlyGlyGly		20	
QY	188	TACAGTAGAAGTGTGAGCAAGTGAACACCGAGGTGAGACGGCAGGTGGTGAATGTCAC	247		
Db	21	-----	Gly-ValValAsnAspAl	26	
QY	248	CTATTCCAACCG--	GAGCAGACCCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTT	304	
		::: :::			
Db	26	aPheLysAspAlaLeuGlnAraGlnileAlaAlaLysileGlyGlyAsp----	44		
QY	305	GGAGAACCATGCTGAAGGTCTCTACATCCCCATGAGCAGATAGACACGGGACCTGA	364		
Db	45	-----	AlaGlyThrSerLe	49	
QY	365	GAATGGGCGCCAGGGGCTTTGGCTCTCGGGGTGAC-----	401		
Db	49	uAsnSerAsnAspTyrGlyTyrGlyGlyGlnLysArgProLeuGluAspGlyAspGlnPr	69		
QY	402	-----	CCCCCGCAGGGCTCAGCTGTGCGGGGGGCCAGGCAAGCA	445	
Db	69	oAspAlaLysLysValProProGlnAsnAspSerPheGlyAlaGlnLeuProProMetHi	89		
QY	446	GCAGCAAGTGGAC-----	ATCCCCCTTCGGCTCTGCTGCGCCACCAGCTATGTGGG	496	
Db	89	sGlnGlnSerArgSerValMetThrGluGluTyrLysValProAspGlyMetValGl	109		
QY	497	TGCCATTATTGGCAAGGAGGGGCCCATCCGCAACATCACAAACAGACCCAGCTCCAA	556		
Db	109	yPheileileGlyArgGlyGlyGluGlnileSerArgilleGlnGlnGluSerGlyCysyl	129		
QY	557	GATAGACTGCATAGGAGGAGACCGAGTGCAGCTGAAGAACCATCAGTGTGCATCT	616		
Db	129	sileGlnle---AlaProAspSerGlyGlyLeuProGluArgSerCysMetLeuThrGl	148		
QY	617	CACCCCTGAGGCTGCTCCTCGCTTGTAAGATGATCTTGGAGATTATGCATAAA----	671		
Db	148	yThrProGluSerValGlnSerAlaLysArgLeuLeuAspGlnIleValGluLysylAr	168		
QY	672	-GAGGCTAAGGACCAAAACGGCTGAC-----	GAGGTTCCCTCGAAGATCTCTGCG	721	
Db	168	gProAlaProGlyPheHisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIl	188		
QY	722	CCATAATTAACCTTTGAGGCGCTCATTTGGCAAGGAGGACGCGAACCTGAAGAGTAGA	781		
Db	198	eProLaserLysAlaGlyLeuValleGlyLysGlyGlyGluThrIleLysGlnLeuGl	208		
QY	782	GCAAGATCCGAGACAAATACCATCTCTCTGTTGCAAGACCTTACCCCTTTCAACCC	841		
Db	208	nGluArgAlaGlyValLysMetValMet-----	IleGlnAsp-----	GlyPr	222
QY	842	TCAGAGGACC-----	ATCACTGTGAAGGGGCCATCGAGATTTGTTGCG	886	
Db	222	oGlnAsnThrGlyAlaAspLysProLeuArgIleThrGlyAspProTyrLysValGlnGl	242		
QY	887	GGCCGAGCAGGAAATAATGAAGAAAGTTCGGGAG-----	GCCTATGAGAA	931	


```

RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May interact with single-stranded DNA from the far-
CC upstream element (FUSE). May activate gene expression.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96I24-1; Sequences=Displayed;
CC Name=2;
CC IsoId=Q96I24-2; Sequence=VSP_008323, VSP_008324;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Detected in a number of cell lines.
CC -!- SIMILARITY: Contains 4 KH domains.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 18.
CC -!- CAUTION: Ref.2 (AAH01325) sequence differs from that shown due to
CC a frameshift in position 527.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U69127; AAC50893.1; ALT_FRAME.
DR EMBL; BC001325; AAH01325.1; ALT_FRAME.
DR EMBL; BC007874; AAH07874.1; -.
DR Genew; HGNC:4005; FUBP3.
DR MIM; 603536; -.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS50084; KH_TYPE_1; 4.
KW Transcription regulation; Trans-acting factor; Nuclear protein;
KW DNA-binding; Repeat; Alternative splicing.
FT DOMAIN 77 141 KH 1.
FT DOMAIN 162 228 KH 2.
FT DOMAIN 253 317 KH 3.
FT DOMAIN 354 421 KH 4.
FT VARSPPLIC 1 64 MAELVQGSAPVGMKAEVDFDLHVRVQTAAKIDIPHLNN
FT STPLVDPSVYGVQKRPDDGV -> MPPI (in
FT isoform 2).
FT /FTId=VSP_008323.
FT /FTId=VSP_008324.
FT Missing (in isoform 2).
FT VARSPLIC 298 548 V -> D (IN REF. 1).
FT CONFLICT 72 72 61640 MW; F1BE223542BC197D CRC64;
FT SEQUENCE 572 AA; 61640 MW; F1BE223542BC197D CRC64;
Alignment Scores:
Pred. No.: 5-59e-09 Length: 572
Score: 262.00 Matches: 126
Percent Similarity: 37.31% Conservatives: 90
Best Local Similarity: 21.76% Mismatches: 218
Query Match: 7.42% Indels: 145
DB: 1 Gaps: 21
US-09-270-437D-7 (1-1946) x FUB3_HUMAN (1-572)
QY 282 ATGAGCTGAATGGC-----CACCAGTTGGAGAACCATGCCCTGAAGATC 326
Db 14 MetLysAlaGluGlyPheValAspAlaLeuHisArgValArgGlnIleAlaAlaIle 33
QY 327 TCCTACATCCCGATGACGAGATAGACACAGGACCTCAGAAATGGCGCCGAGGGGCTTT 386
Db 34 AspSerIleProHisLeuAsnAsnSerThrProLeuValAspSerValTyrGlyTyr 53
QY 387 GGTCTCGGGGTGAGCC---CGCCAGGGGTACCTGTGGCAGCGGGGGCCCGAGCAAG 443
Db 54 GlyValGlnLysArgProLeuAspAspGlyValGlyAsnGlnLeuGlyAlaLeuValHis 73

```

```

QY 444 CAGCAGCAAGTGGACATCCCTCCGCTCTGGTGGCCACCCAGATATGTGGTCCCATTT 503
Db 74 GlnArgThrVal---IleThrGluGluPheLysValProAspLysMetValGlyPheIle 92
QY 504 ATTGGCAGGAGGGGCCACCATCGCAACATCATCAAAACAGACCCAGCTCAAGATAGAC 563
Db 93 IleGlyArgGlyGluGlnIleSerArgIleGlnAlaGluSerGlyCysIleGln 112
QY 564 GTGCATAGGAAGAGACGCGAGGTGCAGCTGAAAAAGCCATCATGTGTCATCCACCCCT 623
Db 113 Ile---AlaSerGluSerGlyIleProGluArgProCysValLeuThrGlyThrPro 131
QY 624 GAGGCTCTCTCTCGCTTGAAGATGATCTTGGAGATTATG----- 665
Db 132 GluSerIleGluGlnAlaLysArgLeuGlyGlnIleValAspArgCysArgAsnGly 151
QY 666 -----CATAAAGAGGCTAAGGACACCAAAACGGCTGACGAGGTTCCCTGGAAGATC 716
Db 152 ProGlyPheHisAsnAspIleAspSerAsnSerThrIleGlnGlu-----Ile 167
QY 717 CTGGCCCAATAAATCTTTTAGGGCGCTCTATGGCAAGGAAGGACGGAACCTGAAGAG 776
Db 168 LeuIleProAlaSerLysValGlyLeuValIleGlyArgGlyGlyGluThrIleLysGln 187
QY 777 GTAGAGCAAGATACCGACACAAATACCATCTCTCGTTGCGAGACCTTACCCTT--- 833
Db 188 LeuGlnGluArgThrGlyValIleValMetValMet-----IleGlnAspGlyProLeuPro 205
QY 834 TACAACCTGTAGAGACCATCATCTGTGAAGGGGGCCATCGAGAATTGTTGAGGGCCGAG 893
Db 206 ThrGlyAlaAspLysProLeuArgIleThrGlyAspAlaPheLysValGlnGlnAlaArg 225
QY 894 CAGGAATAATGAAGAAAGTTCGGAGGCCCTATGAGAAATGATGTGGTCCCATAGCTCT 953
Db 226 GluMetValLeuGluIleIleArgGluLysAspGlnAlaAspPheArgGlyValArgGly 245
QY 954 CACCTGATCCCTGGCTGAACCTGGCTGCTAGCTCTTTCCAGCTCTTCCAGCTCATCCAGCGCA 1013
Db 246 AspPhe----- 247
QY 1014 GTCCCGCGCGCTCCACGACGCGTTACTGGGGCTGCTCCCTATAGCTCTTTATGCGAGCT 1073
Db 248 -----AsnSerArgMetGlyGlySer----- 255
QY 1074 CCCGACGAGGATGTCAGAGTGTATTCGCCCGCCAGGCGAGTGGGGCCCATCATCGGC 1133
Db 256 -----IleGluValSerValProArgPheAlaValGlyIleValIleGly 270
QY 1134 AAGAAGGGCGACCATCAACAGCTCTCCCGGTTTGGCAGCGCCTCCATCAAGATTGCA 1193
Db 271 ArgAsnGlyGluMetIleLysIleGlnAsnAspAlaGlyValArgIleGlnPheLys 290
QY 1194 CCACCCGAA-----ACACCTGACTCCAAAGTTCGTATGTTATCATCATCGACCGCCA 1247
Db 291 ProAspAspGlyIleSerProGlu-----ArgAlaAlaGlnValMetGlyProPro 307
QY 1248 GAG-----GCCCAATTCAAG 1262
Db 308 AspArgCysGlnHisAlaAlaHisIleIleSerGluLeuIleLeuThrAlaGlnGluArg 327
QY 1263 -----CCTCAGGGAAGCAATCATATGCAAACTCAAGGAGGAG 1298
Db 328 AspGlyPheGlyGlyLeuAlaAlaAlaArgGlyArg-----GlyArgGlyArgGlyAsp 345
QY 1299 AACTTCTTTGGTCCCAAGGAGGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCA 1358
Db 346 TrpSerValGlyAlaProGlyGlyValGln---GluIleThrTyrThrValProAlaAsp 364
QY 1359 GCAGCTGCGCGGTCATTGGCAAGGTGGAAGGCGGGAACAGGTGCAAGATTGACG 1418
Db 365 LysCysGlyLeuValIleGlyLysGlyGlyGluAsnIleLysSerIleAsnGlnGlnSer 384

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:51:55 ; Search time 17.1191 Seconds
(without alignments)
11838.037 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtacggagggtggtgg.....atttccttcagggttttaaaa 1946

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xl
-Q=/cgn2_1/USPTO.spool_p/US09270437/runat 23072004 095739 27413/app query.fasta 1.5582
-DB=SwissProt 42 -OPMT=fastan -SUFFIX=rs -MINMATCH=0.1 -IOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1.1 57 @runat 23072004 095739 27413 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	262	7.4	572	1 FUB3 HUMAN	Q96124 homo sapien
2	251	7.1	651	1 FUB1 MOUSE	Q91w78 mus musculus
3	247.5	7.0	643	1 FUB1 HUMAN	Q96ae4 homo sapien
4	244.5	6.9	721	1 FUB2 RAT	Q99pf5 rattus norv
5	237.5	6.7	769	1 ZBP2 CHICK	Q8uvd9 gallus gall
6	236.5	6.7	492	1 NOA2 HUMAN	Q9unw9 homo sapien
7	234.5	6.6	707	1 FUB2 HUMAN	Q92945 homo sapien
8	231	6.5	474	1 NOA1 RAT	Q80wa4 rattus norv
9	221.5	6.3	356	1 PCB1 HUMAN	Q153k6 homo sapien
10	221.5	6.3	493	1 NOA1 MOUSE	Q9jkn6 mus musculus
11	221.5	6.3	510	1 PCB1 HUMAN	P51513 homo sapien
12	218.5	6.2	356	1 PCB1 RABIT	Q19048 oryctolagus
13	218	6.2	1268	1 VGLN HUMAN	Q00341 homo sapien
14	217.5	6.2	403	1 PCB4 HUMAN	P57723 homo sapien
15	216.5	6.1	606	1 Y475 ARATH	P58223 arabidopsis
16	214.5	6.1	403	1 PCB3 HUMAN	P57724 mus musculus
17	208.5	5.9	339	1 PCB3 HUMAN	P57721 homo sapien
18	207.5	5.9	339	1 PCB3_MOUSE	P57722 mus musculus

19	202.5	5.7	1270	1	VGLN_CHICK	P81021 gallus gall
20	199	5.6	365	1	PCB2_HUMAN	Q15366 homo sapien
21	197.5	5.6	463	1	ROK_HUMAN	Q07244 homo sapien
22	197.5	5.6	463	1	ROK_RABIT	O19049 oryctolagus
23	194	5.5	362	1	PCB2_MOUSE	Q61990 mus musculus
24	191	5.4	413	1	PBP2_YEAST	P38151 saccharomyc
25	182.5	5.2	1222	1	S160_YEAST	P06105 saccharomyc
26	178	5.0	381	1	YBD2_YEAST	P38199 saccharomyc
27	170.5	4.8	606	1	TDRH_HUMAN	Q9y2w6 homo sapien
28	164	4.6	784	1	YAV2_XANCV	P14728 xanthomonas
29	156.5	4.4	648	1	GLT0_WHEAT	P10387 triticum ae
30	155.5	4.4	1003	1	MBD6_HUMAN	Q96dne homo sapien
31	154	4.4	839	1	GLT5_WHEAT	P10388 triticum ae
32	153.5	4.3	1820	1	STR0_HUMAN	Q9p2p6 homo sapien
33	153.5	4.3	2716	1	OSA_DROME	Q9in94 drosophila
34	153	4.3	2805	1	MAPA_HUMAN	P78559 homo sapien
35	151	4.3	838	1	GLT4_WHEAT	P08489 triticum ae
36	150	4.2	2142	1	BAT2_HUMAN	P48634 homo sapien
37	147.5	4.2	1461	1	IE18_PRIVIF	P11675 pseudorabie
38	147	4.2	1181	1	HAIR_RAT	P97609 rattus norv
39	145	4.1	660	1	YHL1_EBV	P03181 epstein-bar
40	144	4.1	5147	1	PCLO_HUMAN	Q9y6v0 homo sapien
41	143	4.0	1336	1	W146_HUMAN	Q9c0j8 homo sapien
42	141	4.0	1729	1	TABP_HUMAN	Q9c0c2 homo sapien
43	140.5	4.0	1446	1	IE18_PRIVKA	P33479 pseudorabie
44	140	4.0	1152	1	MAP4_HUMAN	P27816 homo sapien
45	139.5	4.0	931	1	MRTA_HUMAN	Q969v6 homo sapien

ALIGNMENTS

RESULT 1
FUB3_HUMAN STANDARD; PRT; 572 AA.
ID FUB3_HUMAN AC Q96124; Q92946; Q95VB6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Far upstream element binding protein 3 (FUSE binding protein 3).
GN FUBP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97094955; PubMed=8940189;
RA Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;
RT "The far upstream element-binding proteins comprise an ancient family
of single-strand DNA-binding transcription factors.";
RL J. Biol. Chem. 271:31679-31687(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Cervix, and Placenta;
RX MEDLINE=22398257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N.I., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length

This Page Blank (uspto)

[illegible]

C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
R:McKnight, G.L.; Reasoner, J.; Gilbert, T.; Sundquist, K.O.; Hokland, B.; McKernan, P.A.
J. Biol. Chem. 267, 12131-12141, 1992
A:Title: Cloning and expression of a cellular high density lipoprotein-binding protein
A:Reference number: A44125; MUID:92291094; PMID:1318310
A:Accession: A44125
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1268 <MCK>
A:Cross-references: GB:M64098; GB:M83789; NID:g183891; PIDN:AAA35962.1; PID:g183892
A:Note: sequence extracted from NCBI backbone (NCBIN:106862, NCBIIP:106863)
C:Superfamily: vigilin

Alignment Scores:

Pred. No.:	3,99e-07	Length:	1268
Score:	218.00	Matches:	109
Percent Similarity:	38.46%	Conservative:	81
Best Local Similarity:	22.06%	Mismatches:	191
Query Match:	6.17%	Indels:	113
DB:	2	Gaps:	19

US-09-270-437D-7 (1-1946) x A44125 (1-1268)

QY	195	GAGAACTGTGAGCAAGTGAACACCGAGGTGAGACGGCAGTGGTGAATGTCACCTATTCC	254
DB	403	GlugluGluAspLysLeuGluGluGlyProThrGluAspValAsnValAlaGlnGlu	422
QY	255	AACCGGGAGCAG-----ACCAGGCAAGCCATCATGAAGTGAAT---	293
DB	423	GlnIleGluGlyMetValLysAspLeuIleAsnArgMetAspTyrValGluIleAsnIle	442
QY	294	GGCCACAGTGGAGAACCAACATGCCCTG-----AAGTCTCTACATCCCC	338
DB	443	AspHisLysPheHisArgHisLeuIleGlyLysSerGlyAlaAsnIleAsnArgIleLys	462
QY	339	GATGAG-----CAGATAGCACAGGACCTGAGAAATGGGCGCCGAGGGGGCTTTGGC	389
DB	463	AspGlnTyrLysValSerValArgIleProProAspSerGluLysSerAsnLeuIleArg	482
QY	390	TCTCGGGGTGAGCCCGCCAGGGGCTCACCTGTGGCAGCGGGGCCCGCAGCAGACGACG	449
DB	483	IleGluGlyAspPro---GlnGlyVal-----GlnGlnAlaLysArgGlu	496
QY	450	CRAGTGGACATCCCTCTGG-----CCTCTGGTGGCCAC	485
DB	497	LeuLeuGluLeuAlaSerArgMetGluAsnGluArgThrLysAspLeuIleIleGluGln	516
QY	486	CAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATCGCAACATCACAAAACAG	545
DB	517	ArgPheHisArgThrIleIleGlyGlnLysGlyGluArgIleArgGluIleArgAspLys	536
QY	546	ACCACTCCCAAGATA-----GACGTGCATAGGAGGAGAACCGAGGTGCAGCT	593
DB	537	PheProGluValIleIleAsnPheProAspProAlaGlnLysSerAsp-----	552
QY	594	GAAAAAGCCATCAGTGTGACTCCACCCCTGAGGGGTGCTCTCCCTGTTGATAGTATC	653
DB	553	-----IleValGlnLeuArgGlyProLysAsnGluValGluLysCysThrLysTyr	569
QY	654	TTGGAGATTATGCATAAGAGGGCTAAGGACACCAACCAACCGCTACGAGGTTCCCTGAAG	713
DB	570	MetGlnLysMetValAlaAspLeuValGluAsnSerTyrSerIleSerValProIlePhe	589
QY	714	ATCCTGGCCCATATAACTTTGTAGGGCGTCTCATTGGCAAGGAGGACGAACTCTGAAG	773
DB	590	LysGlnPheHisLysAsn-----IleIleGlyLysGlyGlyAlaAsnIleLys	605
QY	774	AAGGTAGACAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTTACCTT	833
DB	606	LysIleArgGluGluSerAsnThrLysIle-----AspLeuProAla	619

QY	834	TACAAACCT---GAGAGGACCATCTGTAAGGGGGCCATCAGAAATGTTTCAGGGCC	890
DB	620	GluAsnSerAsnSerGluThrIleIleThrGlyLysArgAlaAsnCysGluAlaAla	639
QY	891	GAGCAGGAATAATGAAGAAAGTCTGGGAGGCTGATGAATCATGTGGCTGCATGAGC	950
DB	640	ArgSerArgIleLeu-----SerIleGlnLysAspLeuAla-----	651
QY	951	TCTCACTGATCCCTGGCCTGAACCTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGC	1010
DB	652	-----AsnIleAlaGlu-----	655
QY	1011	GCAGTCCCGCCCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGTCTCTTATGCAG	1070
DB	655	-----	655
QY	1071	GCTCCCGAGCAGGATGTGCAGTGTATTATCCCGCCCGCCAGGCAGTGGCGCCATCATC	1130
DB	656	-----ValGluValSerIleProAlaLysLeuHisAsnSerLeuIle	669
QY	1131	GGCAGAGAGGGCAGCAGCATCAACAGCTCTCCGGTTTGGCAGCGCCTCCATCAAGATT	1190
DB	670	GlyThrLysGlyArgLeuIleArgSerIleMetGluGluCysGlyGlyValHisIleHis	689
QY	1191	GCACCACCCGAAACACCTGACTCCAAAGTTCGTATGTTTATCATCTACTGCAGCCAGAG	1250
DB	690	PheProValGluGlySerGlySerAspThr-----ValIleArgGlyProSerSer	707
QY	1251	GCCCAATTCAAGGCTCAGGGAAGAATCTATGCAAACTCAAGGAGGAGAACTCTTTTGGT	1310
DB	708	AspValGluLysAlaLysLysGlnLeuLeu---HisLeuAlaGluGlu-----	722
QY	1311	CCAGAGGAGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGTGGCCGG	1370
DB	723	---LysGlnThrLysSerPheThrValAspIleArgAlaLysProGluTyrHisLysPhe	741
QY	1371	GTCAATGGCAAGGTGGAAAAACGGTGAACGAGTTGCAGAAATTCAGCGCAGCTGAGGTG	1430
DB	742	LeuIleGlyLysGlyGlyLysLysValArgLysValArgAspSerThrGlyAlaArgVal	761
QY	1431	GTAGTACCAAGAGACCCCTGATGAGACGACACAGGTTCATCGTAAAAATCATCGGA	1490
DB	762	IlePhePro-----AlaAlaGluAspLysAspGlnAspLeuIleThrIleGly	778
QY	1491	CATTCTATCCAGTCAGATGCTCAACGGAAGATCCGAGACATCTCTGCCAGGTTAAG	1550
DB	779	LysGluAspAlaValArgGluAlaGlnLysGluLeuGluAlaLeuIle-GlnAsnLeuAs	798
QY	1551	CAGCAGCATCAGAGGGACAGAGTAACAGGCGCCAGGCAC	1590
DB	798	phsValValGluAspSerMetLeuValAspProLysHis	811

RESULT 15
S41224
hnRNP protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C:Accession: S41224
R:Siomi, H.; Matunis, M.J.; Michael, W.M.; Dreyfuss, G.
Nucleic Acids Res. 21, 1193-1198, 1993
A:Title: The pre-mRNA binding K protein contains a novel evolutionarily conserved motif.
A:Reference number: S41224; MUID:93219080; PMID:8464704
A:Accession: S41224
A:Molecule type: mRNA
A:Residues: 1-396 <SIO>
C:Superfamily: transformation upregulated nuclear protein HNRPK

Alignment Scores:			
Pred. No.:	3.08e-07	Length:	396
Score:	217.50	Matches:	90
Percent Similarity:	41.67%	Conservative:	75
Best Local Similarity:	22.73%	Mismatches:	167
Query Match:	6.16%	Indels:	64


```
Db 430 ---SerAlaProAlaGlnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln 448
QY 1662 CTCCTCCGGCAGCCCTGAGAAATGAGTGGGAATCCGGG----- 1697
Db 449 IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTyrGlnGlnAsnSerValTyr 468
QY 1698 -----ACANWGGGCGGGCTGTAGATCAGGTTTG----- 1727
Db 469 AlaGlnGlnThrAlaAlaProAlaAlaProTyrAlaAlaAlaGlyIleValGlnPro 488
QY 1728 -----CCCACTTGATGAGAAAGATGTCCAGTGAGGAACC 1763
Db 489 GlnGlnGlnValAlaTyrGlnGlnProGlnValValGlnGlnVal---GlnProAlaThr 507
QY 1764 CTGATCTNTCAGCCCCCAACACCCACC----- 1790
Db 508 ThrValAlaAlaThrThrThrThrProThrValAspProValThrGlyGlnGlnAspTyrSer 527
QY 1791 ---CAATTGGCCCAACACTGTGTGCCCCCTCGGGGTG----- 1823
Db 528 AlaGlnTrpMetGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 547
QY 1824 -----TCAGAAATNTAGCGCAAGGCACCTTTTAAACGTGGATTG 1862
Db 548 AlaGlnMetLysLysLysLysAlaGluAlaAlaAlaArgAlaValProGlyGlyLeu 567
QY 1863 TTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACCTCAG 1910
Db 568 -----IleGlnGlnMetProMetGlyMetAlaMetProGln 579

RESULT 12
I38489
onconeural ventral antigen-1 - human
C:Species: Homo sapiens (man)
C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 05-Nov-1999
C:Accession: I38489
R:Buckanovich, R.J.; Posner, J.B.; Darnell, R.B.
Neuron 11, 657-672, 1993
A:Title: Nova, the Paraneoplastic Ri Antigen, is Homologous to an RNA-Binding Protein an
A:Reference number: I38489; MUID:94000830; PMID:8398153
A:Accession: I38489
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-510 <RES>
A:Cross-references: EMBL:U04840; NID:G440877; PIDN:AAAL6022.1; PID:G440878
C:Genetics:
A:Gene: Nova-1

Alignment Scores:
Pred. No.: 1.75e-07 Length: 510
Score: 221.50 Matches: 111
Percent Similarity: 34.89% Conservative: 68
Best Local Similarity: 21.64% Mismatches: 173
Query Match: 6.27% Indels: 161
DB: 2 Gaps: 17

US-09-270-437D-7 (1-1946) x I38489 (1-510)
QY 363 GAGATGGGCGCGCAGGGGGCTTT-----GGCTCTCGGGGTCTCAG 401
Db 9 GlnAsnGlyThrHisThrGlyValProIleAspLeuAspProAspSerArgLysArg 28
QY 402 CCCCAGGAGGCTCAGCTGTGGAGCGGGGGCCCGCCAGCAGCAGCAAGTGGAC--- 458
Db 29 ProLeuGluAlaProProGluAlaGlySerThrLysArgThrAsnThrGlyGluAspGly 48
QY 459 ----ATCCCCCTTGGCTCTCGGTGGCCACCAGCATGTGTGGTGCATTATTGGCAGGAG 515
Db 49 GlnTyrPheLeuLysValLeuIleProSerTyrAlaAlaGlySerIleGlyLysGly 68
QY 516 GGGGCCACCATCCGCACATCACAACACAGACCAGTCCCAAGATAGACGTGCATAGAG 575
```

```
Db 69 GlyGlnThrIleValGlnLeuGlnLysGlnThrGlyAlaThrIleLysLeuSerLysLeu 88
QY 576 GAGAACGCA-----GGTCAGCTGAAAAAGCCATCAGTGTGCACCTCCACC 620
Db 89 SerLysSerLysAspPheTyrProGlyThrThrGluArgValCysLeuIleGlnGlyThr 108
QY 621 CTGAGGGCTGCTCTCCGCT-----TCTAAGATGATCTTGGAGATTATGCAT 668
Db 109 ValGluAlaLeuAsnAlaValHisGlyPheIleAlaGlyLysIleArgGluMetProGln 128
QY 669 AAAGAGGCTAAG----- 680
Db 129 AsnValAlaLysThrGluProValSerIleLeuGlnProGlnThrThrValAsnProAsp 148
QY 681 -----GACACCAAAACGGCTGACGAGGTTCC 707
Db 149 ArgIleLysGlnThrLeuProSerSerProThrThrThrLysSerSerProSerAspPro 168
QY 708 CTG-----AAGATCTCGGCCCATATAAATCTTGTAGGGCGT 743
Db 169 MetThrThrSerArgAlaAsnGlnValLysIleIleValProAsnSerThrAlaGlyLeu 188
QY 744 CTCATTGGCAGGAAGGAGCGAACCTGAGAGAGGTAGAGCAAGATACCGAGACAAAAATC 803
Db 189 IleIleGlyLysGlyGlyAlaThrValLysAlaValMetGluGlnSerGlyAlaTrpVal 208
QY 804 ACCATCTCTGCTGTCAGACCTTACCCCTTACACCCCTGAGAGGACCATCTACTGTGAAG 863
Db 209 GlnLeuSerGlnLysProAsp---GlyIleAsnLeuGlnGluArgValValThrValSer 227
QY 864 GGGGCCATCGAATTTGTCAGGCGCGAGCAGGAATAATGAAGAAAGTCTCGGAG--- 920
Db 228 GlyGluProGluGlnAsnArgLysAlaValGluLeuIleGlnLysIleGlnGluAsp 247
QY 921 -----GCCTATGAGAAT----- 932
Db 248 ProGlnSerGlySerCysLeuAsnIleSerTyrAlaAsnValThrGlyProValAlaAsn 267
QY 933 -----CATGTGCTGCCATG 947
Db 268 SerAsnProThrGlySerProTyrAlaAsnThrAlaGluValLeuProThrAlaAla 287
QY 948 AGCTCTCACTCATCCCTGGCCTGAACCTGGCTGTAGTGGTCTTTTCCCA----- 998
Db 288 AlaAlaGlyLeuLeuGlyHisAlaAsnLeuAlaGlyValAlaAlaPheProAlaValLeu 307
QY 998 ----- 998
Db 308 SerGlyPheThrGlyAsnAspLeuValAlaIleThrSerAlaLeuAsnThrLeuAlaSer 327
QY 999 -----GCTTCATCCAGCGCA 1013
Db 328 TyrGlyTyrAsnLeuAsnThrLeuGlyLeuGlyLeuSerGlnAlaAlaThrGlyAla 347
QY 1014 GTCCCCGCCCTCCAGCAGCGTTACTGGGGGTGCTCCCTATAGTCTCTTTATGAGGCT 1073
Db 348 LeuAlaAlaAlaAlaSerAlaAsnProAlaAlaAlaAlaAlaAsnLeuLeuAlaThr 367
QY 1074 CCGGACGAGGAGATGGTGCAGGTGTTATCCCGCCAGCGAGTGGCGGCCCATCATCGGC 1133
Db 368 TyrAlaSerGlu-----AlaSerAlaSerGlySerThrAlaGly 380
QY 1134 AAGAGGGGCGCAGCATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATTGCA 1193
Db 381 GlyThrAla-----GlyThrPheAlaLeuGlySerLeuAlaAla 394
QY 1194 CCACCCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCATCGACCGCCAGAGGC 1253
Db 395 -----ThrAlaAlaThrAsnGly 400
QY 1254 CAATTCAAGGCTCAGGAAAGAAATCTATGCAAACTC-----AAGGAGAGAACTTC 1304
Db 401 TyrPheGlyAlaAlaSerProLeuAlaAlaSerAlaIleLeuGlyThrGlyLysSerThr 420
```

```

Db      408  ---SerAlaProAlaGlnAsnAsnPheGlyGlyGlnGlnGlnTrpAsnProValAlaGln 426
Qy      1662  CTCCTCCGCGCAGGCTGGAATGAGTGGGAATCCGGG----- 1697
Db      427  IleProAlaAlaAlaGlnAsnProTyrGlnValGlyGlyTyrGlnGlnAsnSerValTyr 446
Qy      1698  -----ACACNTGGCGCGGCTGTAGATCAGGTTTG----- 1727
Db      447  AlaGlnGlnThrAlaAlaProAlaAlaAlaProTyrAlaAlaAlaGlyIleValGlnPro 466
Qy      1728  -----CCCACCTTGATTGAGAAAGATGTTCCAGTGGAGAAC 1763
Db      467  GlnGlnGlnValAlaTyrGlnGlnProGlnValValGlnGlnVal---GlnProAlaThr 485
Qy      1764  CTGATCTNTCAGCCGCCAACACCCAC----- 1790
Db      486  ThrValAlaAlaThrThrThrProThrValAspProValThrGlyGluGlnAspTyrSer 505
Qy      1791  ---CAATGGCCCAACACTGNTGCCCTCGGGTG----- 1823
Db      506  AlaGlnTrpMetGluTyrTyrLysSerIleGlyAlaHisAspLysAlaGluAlaValGlu 525
Qy      1824  -----TCAGAAATTNTAGCGCAAGGCACCTTTTAAACGTGGATTG 1862
Db      526  AlaGlnMetLysLysLysLysAlaGluAlaAlaAlaAlaAlaValValProGlyGlyLeu 545
Qy      1863  TTTAAGAAGCTCTCCAGGCCCCCAAGAGGGTGGATCACACCTCAG 1910
Db      546  -----IleGlnGlnMetProMetGlyMetAlaMetProGln 557

RESULT 11
T19217
hypothetical protein C12D8.1b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T19217
R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19092
A;Accession: T19217
A;Molecule type: DNA
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Residues: 1-611 <WIL>
A;Cross-references: EMBL:Z73969; PID:CAA98233.1; GSPDB:GN00023; CESP:C12D8.1b
A;Experimental source: clone C12D8
C;Genetics:
A;Gene: CESP:C12D8.1b
A;Map position: 5
A;Introns: 181/3; 340/3; 535/2

Alignment Scores:
Pred. No.: 3,41e-08 Length: 611
Score: 232.00 Matches: 108
Percent Similarity: 37.15% Conservative: 106
Best Local Similarity: 18.75% Mismatches: 232
Query Match: 6.57% Indels: 130
DB: 2 Gaps: 18

US-09-270-437d-7 (1-1946) x T19217 (1-611)

Qy      435  CCAGCCAGCAGCAGCAGTGGACATCCCGCTT----- 467
Db      50  ProAlaLysArgProAlaAspAspThrAspLeuAsnProPheMetAspAsnGluAla 69
Qy      468  -----CGGTCCCTGGTGCCACCAGTATGTGTGGTGCATTATTGGCAAGGAGGG 518
Db      70  ValAsnGluLysTyrProIleProGluSerAlaValGlyIleValGlyArgGlyGly 89
Qy      519  GCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGATAGGAAGGAG 578
Db      90  SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAspAla 109
```

```

Qy      579  AACCGAGGTGTCAGCTGAAAAAGCCATCATGTGTGACTCCACCCCTGAGGGGTGCTCTCC 638
Db      110  AspProSerSerValArgMetValThrLeuGluGlySerArgSerAsnValGluThr 129
Qy      639  GCTTTGAAGATGATCTTCGAGATTATGCATAAAGAGGCTAAGGACACCAAA----- 689
Db      130  AlalysHisLeuIleAsnGluValValAlaArgSerGlnAsnProArgProGlnTyrGly 149
Qy      690  ---ACGGCTGACGAGGTTCCCTCGAAGATCTCGCCCATATAATACTTTGTAGGGGTCTC 746
Db      150  PheProArgAlaGlnThrThrIleAspIleAlaIleProProAsnArgCysGlyLeuIle 169
Qy      747  ATTGGCAAGAGAGACGCACTGAGAGAGTAGAGCAAGATACCGACACAAAATCACC 806
Db      170  IleGlyLysSerGlyAspThrIleArgGlnLeuGlnLysSerGlyCysLysMetIle 189
Qy      807  ATCTCCTCGTTGCAAGAC---CTTACCCCTTTACACCTCGAGAGGACCATCACTGTGAAG 863
Db      190  Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 207
Qy      864  GGGCCCATCGAGAAATTGTCAGGGCCGAGCAGGAGAAATAATGAAGAAAGTTGGGAGGCC 923
Db      208  GlyAspProGlnLysIleGluLeuAlaLysGlnLeuValAlaGluIleLeuAsnSerGly 227
Qy      924  TATGAGAAATGATGTGGTGCATCAGCTCTCACCTGATCCCTGGCTGAACCTGGCTGCT 983
Db      228  GlyAspGlyAsnGlyGlySer-----GlyLeuGlnMetHis 240
Qy      984  GTAGTCTCTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGCGCTTACTGGG 1043
Db      241  AlaGly-----Gly 243
Qy      1044  GCTGCTCCTATAGTCTCTTTATGCAGGCTCCGAGCAGGAGATGGTGCAGGTGTTTATC 1103
Db      244  GlyGlyGlyAlaSerAla-----ArgGlyGluValValVal 255
Qy      1104  CCCGCCCAGCAGTGGGGCCCATCATCGGCAAGAGGGGAGCAGACATCAACAGCTCTCC 1163
Db      256  ProArgSerValGlyIleIleIleGlyLysGlnGlyAspThrIleLysArgLeuAla 275
Qy      1164  CGGTTTGCAGCGCTCCATCAAGATTGACACCCGCAAAACACCTGACTCCAAAGTTCTGT 1223
Db      276  MetGluThrGlyThrLysIleGlnPheLysProAspAspAspProSerThrProGluArg 295
Qy      1224  ATGTTTATCATCTGAGCCGCCAGAGGCCCAATTCAGAGGCTCAGGGAGAAATTCATGGC 1283
Db      296  CysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGluArgIleThrGlu 315
Qy      1284  AAATCAAGGAGGAG-----AACTTCTTTGGTCCCGAGGAG 1319
Db      316  LeuValLysLysSerThrMetGlnGlnGlyGlyGlyAlaGlyAlaMetVal 335
Qy      1320  GAAGTGAAGCTGGAGACC---CACATAGTGTGCGCAGCATCAGCAGCTGGCGGTCTATT 1376
Db      336  SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 355
Qy      1377  GGCAAGGTGGAAAAACGGTGAACAGAGTTGCAGAAATTTGACGGCAGCTGAGGTGGTAGTA 1436
Db      356  GlyLysGlyGlyGluThrIleLysGlnIleAsnSerGluSerGlyAlaHisCysGluLeu 375
Qy      1437  CCAGAGACACGACCCCTGTATGAGAACGACCGAGTCTATCGTG-----AAA 1481
Db      376  SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyGlyLysArg 395
Qy      1482  ATCATCCGACATTTCTATGCGCAGTCAGATGGCTCAAGGAAGATCCGAGACATCTGGCC 1541
Db      396  AlaIleGluHis-----AlaLysHisLeuIleArgIleLysValGlyAspIleAlaPro 413
Qy      1542  CAGGTTTAAGCAGCAGCATCAGAGGGACAGAGTAAACGAGGCCCGCAGCGGAGGAAGTGA 1601
Db      414  AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe----- 429
Qy      1602  CCAGCCCTCCCTGCTCCCTTGAGTCCAGGACACCAACGGGCAAAATCGAGAGTGTGCT 1661
```



```
Db      422 AlaSerGlyTyrGlyGlnAlaGlyGlnGlyTyrGlnGln-----AspGlyTyrGlyAla 439
QY      1349 GCCAGCATCAGCAGCTGGCCGGGTCAATGGCAAGAGGTGGAAGAAACGGTGAACGAGTTGCA 1408
Db      440 TyrAsnAlaSerGlnGlnSerGlyTyrGlyGln-----AlaAla 452
QY      1409 GAATTTGACGGCAGCTGAGGT---GGTAGTACCAAGACACGACCCCTGTATGAGAACGA 1465
Db      453 GlyTyrAspGlnGlnGlyTyrGlySerThrThrAsnPro-SerGlnGlnGluAsp-- 471
QY      1466 CCAGTCTCATCGTGAATATCATCGGACATCTTATGCCAGTCTGAGTGGCTCAACGGAAGAT 1525
Db      472 -----AlaSerGlnAlaAlaPro----- 477
QY      1526 CCGAGACATCTCGGCCAGGTTTAAGCAGCAGCATCAGAAGGAGCAGAGTAACACGAGCCCA 1585
Db      478 -----ProSerSerAlaGlnSerGlyGlnAlaGlyTyrGlyTh 490
QY      1586 GGCACGGAGGAGTACACGACCCCTCCCTGTCCCTTNGAGTCCAGACACACACGGCGAG 1645
Db      490 rThrGlyGlnGlnProProAlaGlnGlySerThrGlyGlnAlaGlyTyrGlyAlaProPr 510
QY      1646 AAATCGAGAGTGTCTCTCCCGCAGCGCTCAGAAATGAGTGGAAATCCGGGACACNTGG 1705
Db      510 oThrSerGlnAlaGlyTyrSerSerGlnProAlaAlaAlaIarYAsnSerGlyTyrGlyAl 530
QY      1706 GCGGGCTGTAGATCAGGTTTCCCACT 1733
Db      530 aProProAlaSerLysProProThr 539

RESULT 10
T19216
hypothetical protein Cl12D8.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19216
R:McMurray, A.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19092
A:Accession: T19216
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-589 <MIL>
A:Cross-references: EMBL:Z73969; PIDN:CAA98232.1; GSPDB:GN000023; CESP:Cl12D8.1a
A:Experimental source: clone Cl12D8
C:Genetics:
A:Gene: CESP:Cl12D8.1a
A:Map position: 5
A:Introns: 7/3; 25/1; 159/3; 318/3; 513/2

Alignment Scores:
Pred. No.: 3.37e-08 Length: 589
Score: 232.00 Matches: 108
Percent Similarity: 37.15% Conservative: 106
Best Local Similarity: 18.75% Mismatches: 232
Query Match: 6.57% Indels: 130
DB: 2 Gaps: 18

US-09-270-437D-7 (1-1946) x T19216 (1-589)
QY      435 CCAGCCACGACGACGACGAGTGGACATCCCCCTT----- 467
Db      28 ProAlaIysArgProAlaAspThrAspLeuAsnProPheMetAspAsnGluAla 47
QY      468 -----CGGTCCTGGTGGCCACCCAGTATGTGGTCCCATTTATGGCAAGAGGGG 518
Db      48 valAsnGluIysTyrProIleProGluSerAlaValGlyIleValIleGlyArgGlyGly 67
QY      519 GCCACATCCGCAACATCAAAACACACCCAGTCCCAAGATACGACGTGCATAGCAAGAG 578
Db      68 SerGluIleGlnGlyIleGlnAlaLysAlaGlyCysArgValGlnMetSerProAla 87
```

```
QY      579 AACGAGGTGTCAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGCTGCTCTCC 638
Db      88 AspProSerSerGlyValArgMetValThrLeuGluGlySerArgSerAsnValGluThr 107
QY      639 GCTTGTAGATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACACAAA----- 689
Db      108 AlalysHisLeuIleAenGluValAlaAlaArgSerGlnAsnProArgProGlnTyrGly 127
QY      690 ---ACGGCTGACGAGGTCCCTCCTGAAAGTCTCTGGCCCAATAATACTTTGTAGGGGTCTC 746
Db      128 PheProArgAlaGlnThrThrIleAspIleAlaIleProAsnArgCysGlyLeuIle 147
QY      747 ATTGGCAGGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAATACCC 806
Db      148 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnGlyLysSerGlyCysLysMetIle 167
QY      807 ATCTCTCTCTGTGCAGAC---CTTACCTCTTTACACCTCTGAGAGGACCATCTCTGAAG 863
Db      168 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 185
QY      864 GGGGCCATCGAATATTGTTGAGGGCCGAGCAGGAAATAATGAAGAAAGTTTCGGAGGCC 923
Db      186 GlyAspProGlnLysIleGluLeuAlaLysGlnLeuValAlaGluIleLeuAsnSerGly 205
QY      924 TATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGATCCCTGGCTGAACCTGCTGCT 983
Db      206 GlyAspGlyAsnGlyGlySer-----GlyLeuGlnMethHis 218
QY      984 GTAGTCTCTTTCCACAGCTTCATCCAGCAGCTCCCGCGCTCCACGAGCGGTTACTGGG 1043
Db      219 AlaGly-----Gly 221
QY      1044 GCTGCTCCCTATAGTCTCTTTATGCGAGCTCCGAGCAGGAGATGGTCAGGTGTTATC 1103
Db      222 GlyGlyGlyAlaSerAla-----ArgGlyGluValValVal 233
QY      1104 CCGCGCCAGGAGTGGCGGCCATCATCGCAAGAGGGGAGCAGACATCAACAGCTCTCC 1163
Db      234 ProArgSerSerValGlyIleIleIleGlyLysGlnGlyAspThrIleLysArgLeuAla 253
QY      1164 CGGTTTGGCAGCGCTCCATCAAGATTGCACCCGAAACACCTGACTCCAAAGTTGCT 1223
Db      254 MetGluThrGlyThrLysIleGlnPheLysProAspAspProSerThrProGluArg 273
QY      1224 ATGGTTATCATCTGCGACCGCCAGGCCCCAATTCAGGCTCAGGAGATCTATGCG 1283
Db      274 CysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGluArgIleThrGlu 293
QY      1284 AAACCTCAAGGAGGAG-----AACTTCTTGGTCCCAAGGAG 1319
Db      294 LeuValLysLysSerThrMetGlnGlnGlyGlyGlyAlaAlaGlyAlaMetVal 313
QY      1320 GAAGTGAAGCTGGAGACC---CACATACGTGTGCCAGCATCAGACGTGGCGGGTCATT 1376
Db      314 SerAsnGluAlaSerThrPheTyrMetSerValProAlaAlaLysCysGlyLeuValIle 333
QY      1377 GCGAAAGTGGAAACCGTGAACGAGTTGCAGATTTCACGCGCAGCTGAGGTGGTAGTA 1436
Db      334 GlyLysGlyGlyGluThrIleLysGlnIleAsnSerGluSerGlyAlaHisCysGluLeu 353
QY      1437 CCAAGAGACGACACCCCTGATGAGAACACACAGGTCATCGTG-----AAA 1481
Db      354 SerArgAspProThrGlyAsnAlaAspGluLysValPheValIleLysGlyGlyLysArg 373
QY      1482 ATCATCGGCATTTCTTATGCCAGTCCAGTGTGTCAACGGAAGATCCGACATCTCTGGCC 1541
Db      374 AlaIleGluHis-----AlalysHisLeuIleArgIleLysValGlyAspIleAlaPro 391
QY      1542 CAGGTTAAGCAGCAGCATCAGAGGACAGAGTAACACGCCCGCCAGCGAGGAGTGA 1601
Db      392 AsnThrProPheArgAspAspSerAlaMetThrMetGlnThrGlnPhe----- 407
QY      1602 CCAGCCCTCCCTGTCTCCCTTNGAGTCCAGGACAAACACCGGAGAAATCGAGAGTGTGCT 1661
```



```
QY 681 -----GACACAAAGCGCTGACGAGGTCCCTGGAAGATCTCTGGCCCATTAATAC 731
Db 233 ProtleProAlaValSerAsnSerLysValThrIlelleProIleProAlaAsn 252
QY 732 TTTGTAGGCGCTCATTTGGCAAGGAGGACGACCTGGAAGAGGTAGACGATACC 791
Db 253 LysCysGlyAlaIlelleGlyLysGlyGluGlnMetArgLysLeuArgSerTrpThr 272
QY 792 GAGACAAAATACCATCTCTCGTTGCAAGACCTTACCTTTACACCCCTGAGAGACC 851
Db 273 AsnCysAspPheIleLeu-----IleGlnGluAsnAsnIleAlaAspSerValLysPro 290
QY 852 ATCACTGTGAAGGGGCCATGACGAAATTTGTCAGGGCCGAGCAGGAATATGAGAAA 911
Db 291 LeuGlnIleThrGlyGln-----ProLysGlu 299
QY 912 GTTCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCTCTACCTGATCCTCGGCTG 971
Db 300 ValGluHisAla-----LysAlaLeuValAlaAspIleLeuAspGlyPhe 314
QY 972 AACCTGGCTGTAGGTCTTTTCCAGGCTTCATCCAGCGAGTCCCGCCCTCCACG 1031
Db 315 AspGluCys-----ProProAla 320
QY 1032 AGCGTTACTGGGCTGCTCCCTATAGCTCTTATGCAAGCTCCCGCAGCAGAGATGCTG 1091
Db 321 GlyMetAlaGlyAsnSerProValAlaAlaMet-----SerLeu 333
QY 1092 CAGGTGTTTATCCCGCCGAGGAGTGGCGGCATCATCGCAAGAGGGGAGCAGCATC 1151
Db 334 GlnValLysValProArgSerThrValGlyAlaIleMetGlyLeuGlnGlySerAsnIle 353
QY 1152 AAACAGCTCTCCGGTTGCGCGCCTCCATCAAGATGTCACACCCGAAACACTGAC 1211
Db 354 LysLysIleSerAsnGluThrGluThrLysIleGlnPheMetProAspAspProLys 373
QY 1212 TCCAAAGTTCGTATGGTATCATCATCTGACCGCCAGCGAGCCCAATTCAGGCTCAGGA 1271
Db 374 LeuMetGluArgThrLeuValValIleGlyAsnLysAsnLysValTyValCys---Ala 392
QY 1272 AGAATCTATGGCAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAGTGAAGCTG 1331
Db 393 ArgLeuLeuGlnIleValGluAlaAsnSerGluAsnAlaAsnThrProIleSerLeu 412
QY 1332 GAGACCCATACGTGTGCGCAGCATCAGCAGCTGGCCGGGTCTATTGGCAAGAGTGA 1391
Db 413 ---PheTyrMetLeuIleProAlaSerLysCysGlyLeuValIleGlyArgGlyGly 431
QY 1392 ACGGTCAAGAGCTTCAGAAATTTGACGCGAGCTGAGTGTGTAGTACCAAGAGAC 1445
Db 432 ThrIleArgGlnIleAsnLysGluSerGlyAlaTyCysGluMetSerArgAsp 449

RESULT 8
T49962
Hypothetical protein F8M21.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49962
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24493
A/Accession: T49962
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-568 <BEV>
A/Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160
A/Experimental source: cultivar Columbia; BAC clone F8M21
C/Genetics:
A/Gene: ATSP:F8M21.160
A/Map position: 5
A/Introns: 200/3; 337/3; 544/3
```

```
Alignment Scores:
Pred. No.: 2.23e-08 Length: 568
Score: 234.50 Matches: 92
Percent Similarity: 42.47% Conservative: 97
Best Local Similarity: 20.67% Mismatches: 175
Query Match: 6.64% Indels: 81
DB: 2 Gaps: 16

US-09-270-437d-7 (1-1946) x T49962 (1-568)
QY 336 CCAGATGAGCAGATAGCAGGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGG 395
Db 3 ProAspHisArgMetSerProAspHisArgAspSerHisArg-----LysArg 18
QY 396 GGTCAAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCCAGCAGCAG 449
Db 19 SerArgProGlnSerAspTyAspAsnGlyGlySerLysArgArgTyArgGlyAsp 38
QY 450 -----CMAGTGACATCCCGCTTCGCTCCTCGTGGCCACC 485
Db 39 AspArgAspSerLeuValIleAspArgAspAspThrValPheArgTyLeuCysProVal 58
QY 486 CAGTATGTGGTGGCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAACAG 545
Db 59 LysLysIleGlySerValIleGlyArgGlyGlyAspIleValLysGlnLeuArgAsnAsp 78
QY 546 ACCAGTCCAGATAGACGTGCATAGGAGGAGGAGCAGGTGCAGCTGAAAGGCCATC 605
Db 79 ThrArgSerLysIleArgIle---GlyGluAlaIleProGlyCysAspGluArgValIle 97
QY 606 AGTGTGCACTCCACCCCT-----GAGGGCTGCTCTCTCC 638
Db 98 ThrIleTySerProSerAspGluThrAsnAlaPheGlyAspGlyGluLysValLeuSer 117
QY 639 GCTGTAGATGATCTTCGAGATATGATTAAGAG-----GCTAAGACACCAAAACG 692
Db 118 ProAlaGlnAspAlaLeuPheArgIleHisAspArgValAlaAspAlaArgSer 137
QY 693 GCTGAC-----GAGGTCCCTGGAAGATCTCGGCCCATTAATACCTT 734
Db 138 GluAspSerProGluGlyGluValThrAlaLysLeuLeuValProSerAspGln 157
QY 735 GTAGGCGCTCTCATTTGGCAAGGAGGACCGAACCTTGAAGAAGGTAGACAGATCCGAG 794
Db 158 IleGlyCysIleLeuGlyArgGlyGlyGlnIleValGlnAsnIleArgSerGluThrGly 177
QY 795 ACAAATACCATCTCTCTGTTGCAAGACCTTACCTT-----TACAACCTCAGAGG 848
Db 178 AlaGlnIleArgIleValLysAspArgAsnMetProLeuCysAlaLeuAsnSerAspGlu 197
QY 849 ACCATCACTGTGAAGGGGCCATCGAGAATTTGTCAGGGCCGAGCAGGAAATATGAAG 908
Db 198 LeuIleGlnIleSerGlyGluValleuIleValLysLysAlaLeuGlnIleAlaSer 217
QY 909 AAAGTTCGGAGGCGCTATGAGAATCATGTGGCTGCCATGAGCTCTACCTGATCCTCGC 968
Db 218 ArgLeu-----HisGluAsn-----ProSerArgSerGlnAsnLeu----- 229
QY 969 CTGAACCTGGCTGTGTAGTCTTTTCCAGGCTTCATCCAGCGCAGTCCCGCGCTCC 1028
Db 230 -----LeuSerSerSerGlyGlyTyProAlaGlySerLeuMetSerHisAlaGlyGly 247
QY 1029 AGCAGGTTACTGGGGCTGCTCCCTATAGCTCC----- 1061
Db 248 ProArgLeuValGlyLeuAlaProLeuMetGlySerTyGlyArgAspAlaGlyAspTrp 267
QY 1062 -----TTTATGCAGGCTCCCGCAGCAGGAG-----ATGCTGCAG 1094
Db 268 SerArgProLeuTyGlnProProArgAsnAspProProAlaThrGluPheIleArg 287
QY 1095 GTGTTTATCCCGCCCGCAGCAGTGGGGCCATCATCGGCAAGAGGGGAGCAGCATCANA 1154
Db 288 LeuValSerProValGluAsnIleAlaSerValIleGlyGlyAlaLeuIleAsn 307
```

QY 861 AAGGGCCATCGAATTGTTGACGGCCGACGAGGAATAATGAAGAAGTTCCGGAG 920
Db 238 ThrGlyAspProTyrLysValGlnGlnAlaLysGluMetValLeuGluLeuLeuArgAsp 257
QY 921 -----GCATAGAAATGATGTGGCTGCCATGAGCTCTCACCTGATCCCT 965
Db 258 GlnGlyGlyPheArgGluValArgAsnGluTyrGly ----- 269
QY 966 GGCGTGAACCTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCT 1025
Db 269 ----- 269
QY 1026 CCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCCCTTATGAGGCTCCGACGAGAG 1085
Db 270 ---SerArgIleGlyGly-----AsnGlu 276
QY 1086 ATGTGTGAGGTGTTTATCCCGCCAGCGAGTGGCGCATCATCGCAAGAGGGCGAG 1145
Db 277 GlyIleAspValProIleProArgPheAlaValGlyLeuValIleGlyArgAsnGlyGlu 296
QY 1146 CAGATCAACAGCTCTCCGGTTTGCAGCGCTCCATCAAGATTGACCAACCCGAA --- 1202
Db 297 MetIleLysLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspAspGly 316
QY 1203 ---ACACCTGACTCCAAAGTTGCTATGCTTATCATCTGACCGCCAGAG---GCCCAA 1256
Db 317 ThrThrProGlu-----ArgIleAlaGlnIleThrGlyProProAspArgCysGln 333
QY 1257 TTCAAGGCTCAG----- 1268
Db 334 HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353
QY 1269 -----CGAAGAANTTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAG 1316
Db 354 ProGlyProGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro 373
QY 1317 GAGAAAGTGAAGCTGAGACCCACATAGCTGTGCCAGCATCAGCAGCTGCCGGTCAAT 1376
Db 374 GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIle 392
QY 1377 GCGAAAGTGGAAAAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTA 1436
Db 393 GlyLysGlyGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeu 412
QY 1437 CCAAGACACAGACCCCTGTATGAGAACGACCAAGTGC---ATCGTGAATAATCATCGACAT 1493
Db 413 GlnArgAsnProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432
QY 1494 -----TTCTATGCCAGTCAGATGGCTCAACCGAAGAT-----CCGAGACAT 1534
Db 433 ProGlnGlnIleAspTyrAlaArgGlnLeuIleGluLysIleGlyGlyProValAsn 452
QY 1535 CTT---GGCCCGAGTTAAAGCAGCAGCATCAGAAGGACAGAGTAACAGAGCCCGAGGACG 1591
Db 453 ProLeuGlyProProValProHisGlyProHisGlyVal-----ProGlyPro---HisG 470
QY 1592 GA-----GGAAGTGAACAGCCCTC-----CTGTCCTCTTNGA 1624
Db 470 LysProGlyProProGlyProGlyThrProMetGlyProTyrAsnProAlaProTyrA 490
QY 1625 GTCACAGA----- 1632
Db 490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510
QY 1633 -----CAACAACGGGCAGAAATCGAGAGTGTCTCTCCCGCG 1669
Db 510 xPglyAsnAlaTyrProHisTrpGlnGlnAlaProProAsp-----ProA 526
QY 1670 CAGGCTGGAATGATGGGAATCCGGGACACWNTGGCGGGCTGTAGATCAGGTTGCC 1729
Db 526 IalysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyr-----TyrAla 542
QY 1730 CACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTNTCAGCCCCCAACCCAC 1789

Db 543 HisTyrTyrGlnGlnGlnAla-----GlnProProProAlaAla 555
QY 1790 CCAATTGGC---CCAACACTGTTNTGCCCTCGGGTGTCAAAAATTTTAGCGCAAGGCAC 1846
Db 556 ProAlaGlyAlaProThrThrGlnThrAsnGlyGlnGlyAsp----- 570
QY 1847 TTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGCGCCCAAGAGAGGGTGTGATCACACC 1906
Db 571 -----GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583
QY 1907 -----TCAGTGGGAAGAAAATAAATTTCTCTTCAGGT 1939
Db 584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602
RESULT 7
T25832
hypothetical protein M01A10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T25832
R:Sheet, P.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid M01A10.
A:Reference number: Z20094
A/Accession: T25832
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-680 <SCH>
A/Cross-references: EMBL:U88174; PIDN:AAB42272.1; GSPDB:GN00019; CESP:M01A10.1
A/Experimental source: strain Bristol N2; clone M01A10
C:Genetics:
A:Gene: CESP:M01A10.1
A:Map position: 1
A:Introns: 63/2; 97/3; 121/1; 160/1; 269/3; 411/3; 512/1; 649/2
Alignment Scores:
Pred. No.: 8,98e-09 Length: 680
Score: 240.50 Matches: 82
Percent Similarity: 41.21% Conservative: 82
Best Local Similarity: 20.60% Mismatches: 159
Query Match: 6.81% Indels: 75
DB: 2 Gaps: 12
US-09-270-437D-7 (1-1946) x T25832 (1-680)
QY 342 GACAGATGACACAGGACCTGAGATGGCGCGAGGGGCTTGGCTCTCGGGTCAG 401
Db 97 LysHisIleSerSerSerThrAspThrAlaArgLysArgAspPheAspGluArgSer--- 115
QY 402 CCGCGCAGCGGCTCACCT-----GTGGCAGCGGGGGCCCGCAAGCAGCAGCAAGTG 455
Db 116 -----GluGlySerAspGluTyrGluGluTyrAlaProProCysLysLeuThrLysGly 133
QY 456 GACATCCCTTCGG-----CTCTGGTGGCC 482
Db 134 AspIleAspTyrArgValAspThrSerThrThrValIleLysAlaSerValSerIlePro 153
QY 483 ACCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCACAAA 542
Db 154 GluGluSerValGlyLeuValIleGlyArgAsnGlyValGluIleGlnAlaIleSerGln 173
QY 543 CAGACCCAGCTCCAAAGATAGACGTGCATAGGAAGAGAAACGAGGTGCAGCTGAAAAAGCC 602
Db 174 LysSerGlyCysArgValGlnIle---ValAlaGluProSerThrThrGlyTyrArgSer 192
QY 603 ATCAGTGTGACATCCACCCCTGAGGGCTGCTCTCGCTTGTGAAGATGATCTTGGAGATT 662
Db 193 ValAspIleTyrGlyLysSerGluAsnIleGluValAlaLysLysLeuIleAsnGluVal 212
QY 663 ATCATTAAGAGCGCTAAG----- 680
Db 213 ValAlaArgGlyArgLysLeuSerGlnGluProLeuProCysSerValProGlnPheGln 232

```
Db 102 HisPheGlyAspGlyGluLeuValCysProAlaLeuAspAlaLeuPheLysValHis 121
QY 657 GAGATTATGCATAAAGAGGCT---AAGGACACCAAAACGGCTGACGAG----- 701
Db 122 AspMetValValAlaAspAlaAspGlnAspAspGlyThrAspAspAsnAspLeuGly 141
QY 702 -----GTTCCCTCTGAGATCTCGGCCCATATACTTTGTAGGGCGTCTATT 749
Db 142 GluLysGlnThrValThrValArgMetLeuValProSerAspGlnIleGlyCysValIle 161
QY 750 GCGAAGGAGGACGACGCTCAAGAGGTAGCAGCAATCCGAGACAAAATCACCATC 809
Db 162 GlyLysGlyGlnValIleGlnAsnLeuArgAsnAspThrAsnAlaGlnIleArgVal 181
QY 810 TCCTCGTTGCAAGAC-----CTTACCCTTTACAACTGAGGAGGACC 851
Db 182 -----IleLysAspHisLeuProAlaCysAlaLeuThrLeuSerHisAspGluLeuLeu 199
QY 852 ATCACTGTGAAGGGGGCCATCGAGAAATGTTCAGGGCCGAGCAGGAAATATGNAGAAA 911
Db 200 LeuIleIleGlyGluProLeu-----Val 207
QY 912 GTTCGGGAGGCC---TATGAG-----AATGATGTGCTGCCATGAGCTCT 953
Db 208 ValArgGluAlaLeuTyrGlnValAlaSerLeuLeuHisAspAsnProSerArgPheGln 227
QY 954 CACCTGATCCCTGGCTGAACCTGGCTGCTGTGTAGCTTTTCCAGCTTCATCCAGCGCA 1013
Db 228 HisLeu-----LeuLeuSerSerSerSerSer 237
QY 1014 GTCCCGCGCCCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTCTCTTTATGCG--- 1070
Db 238 MethHisGlnProGlyAlaMetLeuMetSerAlaAlaLeuThrSerSerHisArgAsnTyr 257
QY 1071 -----GCTCCGAGCAGAGATGGTGCAGGTGTTATC---CCCCGC 1109
Db 258 AlaValArgArgAspIleAlaAspAlaArgGluPheCysValCysPheIleCysProAla 277
QY 1110 CAGGCAGTGGGCGCCATCATCGCAAGAGAGGGCGCAGCACATCAACAGCTCTCCCGGTTT 1169
Db 278 GluAsnValGlyGlyValIleGlyLysGlyGlyPheIleAsnGlnIleArgGlnGlu 297
QY 1170 GCCAGGCTCTCAAGATTGCACCCCGCAACCTGACTCCCAAGATTCTGTATGTT 1229
Db 298 ThrGlyAlaThrIleArgValAsnThrSerGluThrAspAspAsp-----CysIle 315
QY 1230 ATCATCACTGGACCGCCAGAG-----GCCCAATTTC 1259
Db 316 IlePheIleSerSerLysGluPheTyrGluAspGlnSerProAlaValAsnAlaIle 335
QY 1260 AAGGCTCAGGGAAGAAATCTATGGCAACTC---AAGGAGGAGAACTTTCTTGTCTCCAG 1316
Db 336 ArgLeuGlnGlnArgCysSerGluLysValGlyLysAspAlaAsn----- 350
QY 1317 GAGGAAGTGAAGTGGAGCCACATACCTGTGCCAGCATCAGCAGCTGCGCGGTTCATT 1376
Db 351 ---AspLeuAlaIleSerThrArgLeuLeuValSerSerSerGlnIleGlyCysLeuIle 369
QY 1377 GGCAAAGGTGGAAAACCGGTGACAGATTTCAGCGGACGCTGAGGTG---GTA 1433
Db 370 GlyLysGlyGlyAlaValIleSerGluMetArgSerValThrArgAlaAsnIleArgIle 389
QY 1434 GTACCAAGAGACGACCCCTGAT-----GAGAACGACAGGTGATCGTGAATAATCATC 1487
Db 390 LeuGlnLysGluAspValProLysIleAlaArgGluAspGluMetValGlnIleThr 409
QY 1488 GGCATTTCTATCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTCGGCCAGGTT 1547
Db 410 Gly-----SerProAspAlaAlaMetLysAlaLeuThrGlnValIleLeuArgLeu 426
QY 1548 AAG 1550
Db 427 Arg 427
```

RESULT 6

```
A53184
myc far upstream element-binding protein - human
N:Alternate names: FUSE-binding protein
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: A53184
R:Duncan, R.; Bazar, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan, M.; Levens,
Genes Dev. 8, 465-480, 1994
A:Title: A sequence-specific, single-strand binding protein activates the far upstream e.
A:Reference number: A53184; MUID:94170991; PMID:8125259
A:Accession: A53184
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <DUN>
A:Cross-references: GB:U05040; NID:g460151; PIDN:AAAL7976.1; PID:g460152
C:Keywords: DNA binding

Alignment Scores:
Pred. No.: 5.04e-09 Length: 644
Score: 244.00 Matches: 148
Percent Similarity: 36.46% Conservative: 93
Best Local Similarity: 22.39% Mismatches: 219
Query Match: 6.91% Indels: 202
DB: 2 Gaps: 33

US-09-270-437D-7 (1-1946) x A53184 (1-644)
QY 309 AACCATGCCCTGAAGGTCTCTACATCCCGCATGAGCAGATAGCACAG-----GGACCT 362
Db 28 AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGlyGly 47
QY 363 GAGAAATGGG-----CGCCGAGGGGCTTGGCTCTCGGGTGCAGCCCGCCAGCGGC 413
Db 48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyGlnLysArg----- 65
QY 414 TCACCTGNGCGACGGGG----- 431
Db 66 ---ProLeuGluAspGlyGlnProAspAlaLysLysValAlaProGlnAsnAspSer 84
QY 432 -----GCCCCAGCAAGCAGCAGCAGTGGACATCCCTCCCTT-----CGG 470
Db 85 PheGlyThrGlnLeuProProMetHisGlnGlnSerArgSerValMetThrGluGlu 104
QY 471 CTCCTGGTGGCCACCAGTATGTGGTGCATTTATGSCAAGGAGGGGCCACCATCCGC 530
Db 105 TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGlnIleSer 124
QY 531 AACATCACAAACACACCCAGTCCCAAGATAGACGTGCATAGGAGGAGAACCGAGTGCA 590
Db 125 ArgIleGlnGlnGluSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeu 143
QY 591 GCTGAAAAGCCATCAGTGTGCACCTCCACCTGAGGGGTGCTCTCCGCTTGTAGATG 650
Db 144 ProGluArgSerCysMetLeuThrGlyThrProGlnSerValGlnSerAlaLysArgLeu 163
QY 651 ATCTTGGAGATTATGCATAAA-----GAGGCTAAGGACACCAAAACGGCTGAC----- 698
Db 164 LeuAspGlnIleValGluLysGlyArgProAlaProGlyPheHisGlyAspGlyPro 183
QY 699 ---GAGGTTCCCTCCTGAGATCCTGCCCATATAACTTTGTAGGGCGTCTATTGGCAG 755
Db 184 GlyAsnAlaValGlnGluIleMetIleProAlaSerLysAlaGlyLeuValIleGlyLys 203
QY 756 GAAGGACGGAACCTGAAGAAGGTAGCAAGATACCGAGACAAAATCACCATCTCCTCG 815
Db 204 GlyGlyGluThrIleLysGlnLeuGlnArgAlaGlyValLysMetValMet----- 221
QY 816 TTGCAAGACCTTACCCTTTACACCTGAGAGGACC-----ATCACTGTG 860
Db 222 IleGlnAsp-----GlyProGlnAsnThrGlyAlaAspLysProLeuArgIle 237
```



```
QY 438 -----GCCACAGCAGCAGCAAGTGGACATCCCTCTGGCTCTGGTGCACCCAGCATAT 491
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 ThrTyrAlaThrGlnGln-----LeuThrLeuArgAlaLeuLeuSerThrArgGlu 104
QY 492 GTGGGTGCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAG 551
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
105 AlaGlyIleIleIleGlyAlaGlyLysAsnValAlaGluLeuArgSerThrThrAsn 124
QY 552 TCCAAGATAGAGCTGATAGAGGAGGAGACCGAGGTGCAGCTGCAAAAACCCATCAGTGTG 611
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
125 ValLysAlaGlyVal---ThrLysAlaValProAsnValHisAspArgValLeuThrIle 143
QY 612 CACTCACCCCTGAGGGCTGCTCCCTGCTTGAAGATGATCTTGAGATTATGATATAA 671
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 SerGlyProLeuGluAsnValValArgAlaTyrArgPheIleAspIlePhe----- 161
QY 672 GAGGCTAAGCAGACCAAAACGGCTGACGAGGTTCC----- 707
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
162 ---AlaLysAsnSerThrAsnProAspGlyThrProSerAspAlaAsnThrProArgLys 180
QY 708 CTGAAGATCTGGCCCATATAATTTGTAGGGCGTCTCATTTGGCAAGGAGGACCGAAC 767
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 LeuArgLeuLeuIleAlaHisSerLeuMetGlySerIleIleGlyArgAsnGlyLeuArg 200
QY 768 CTGAAGAAGTAGACAGATACCGAGACAAATAATCACCATCTCTCTGTTGCAAGACCTT 827
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 IleLysLeuIleGlnAspLysCysSerCysArgMetIleAlaSer-----LysAspMet 218
QY 828 ACCCTTTACACCTCAGAGGACCATCACTGTGAAGGGGCCATCAGAGAAATGTTGCAGG 887
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
219 LeuProGlnSerThrGluArgThrValGluIleHisGlyThrValAspAsnLeuHisAla 238
QY 888 GCCGAGCAGGAATAATGAAGAAAGTTCGGAGGCTCATGAG----- 929
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 AlaIleTyrGluIleGlyLysCysLeuIleAspAspTyrGluArgGlyAlaGlyThrVal 258
QY 930 -----AATGATGTGCTCCATGAGCTCTCACCTGATCCCTGGCTGAACCTGGCTGT 983
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
259 PheTyrAsnProValSerArgLeuThrGln-----ProLeuProSerLeuAlaSer 275
QY 984 GTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCT-----CCCAGCAGCGTT 1037
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
276 Thr-----AlaSerProGlnGlnValSerProProAlaAlaProSerThrThr 291
QY 1038 ACTGGGGCTGCT-----CCCTATAGTCTC----- 1061
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
292 SerGlyGluAlaIleProGluAsnPheValSerTyrGlyAlaGlnValPheProAlaThr 311
QY 1062 -----TTATGAGGCTCCGAGCAGGAGATGTCAGGTGTTATCCCGCCAG 1112
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
312 GlnMetProPheLeuGlnGlnProLysValThrGln---AsnIleSerIleProAlaAsp 330
QY 1113 GCAGTGGGGCCCATCATCGCAAGAGGAGGAGCAGACATCAAAACAGCTCTCCCGTTTGC 1172
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
331 MetValGlyCysIleIleGlyArgGlySerLysIleSerGluIleArgArgThrSer 350
QY 1173 AGGCGCTCATCAAGATTGCACACCGCAACACCTGACTCCAAAGTTTGTATGGTTATC 1232
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
351 GlySerLysIleSerIleAlaLysGluProHisAspGluThrGlyGluArgMetPheThr 370
QY 1233 ATCACTGGACCGCCAGGCCCCAATTCAGGCTCAGGAGAGATCTATGCAAACTCAAG 1292
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
371 IleThrGlyThrHisGluGluAsnGluLysAlaLeuPheLeuLeuTyrGlnGlnLeuGlu 390
QY 1293 GAGGAG 1298
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 MetGlu 392
```

RESULT 4

T27855

hypothesized protein Zk418.9 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T27855
R;Fulton, L.
submitted to the EMBL Data Library, April 1994
A;Description: The sequence of C. elegans cosmid ZK418.
A;Reference number: Z20430
A;Accession: T27855
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-768 <FUL>
A;Cross-references: EMBL:U00047; PIDN:AAA50693.1; CESP:ZK418.9
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:ZK418.9
A;Introns: 166/1; 221/3; 243/1; 279/1; 387/3; 490/1; 567/3; 727/2

Alignment Scores:

Pred. No.:	9,06e-10	Length:	768
Score:	255.00	Matches:	154
Percent Similarity:	37.13%	Conservative:	110
Best Local Similarity:	21.68%	Mismatches:	292
Query Match:	7.22%	Indels:	155
DB:	2	Gaps:	26

US-09-270-437D-7 (1-1946) x T27855 (1-768)

```
QY 23 GTGCTCTGTCCCTTCCTTCGGCGCTGGCGCTCAGCCACCAGAGCCGGGTGGGA 82
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
95 AlaIlePheSerThrSerMetArgValPheAsnGluLysLysSerTyrMetHisPhePro 114
QY 83 GGGCGAGTGCTCAGCTTCCCGGTTAGGACGGGAAATTCAAATCCGAAATATTCCACC 142
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
115 GlyAsnValAspIlePheProLuhisSerSerGlnLeuPheAsnProHisPheLeuAsn 134
QY 143 CCAGTCCGATGGGAAGTACTGCACAGCTGCTGCTCAGTATGTTACAGTAGAAC-- 200
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 -LysIleArgSerGlnLeuIleGlyLysLeuIle---GluLeuAspIleThrLysAsnG1 153
QY 201 -----TGTAGCAAGTGAACACCGAGAGTGAGACGGC 232
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
153 uSerLeuLeuLeuSerAlaLeuPheLeuCysSerAlaValHisProAspIleSerThrG1 173
QY 233 AGTGTGTGATGTCCTCTATTCACCGGAGCAGACCGAGCAAGCATCATG-- 284
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
173 uGlyLysAsnLeuLeuTyrLysTyrGlnGlnTyrSerAlaAlaLeuIleAsnHisCy 193
QY 285 -----AAGCTGAATGGCCACCATGTTGGAGAACCATGCCCTGAAGTCTCTCA 331
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 sCysLeuThrAsnGlnAsnAlaProSerArgTyrSerGluLeuLeuSerLeuTyrG1 213
QY 332 CATCCCGATGACAGATAGACAGGACCTGAGATGGCGCGCGCGGGGCTTTGGCTC 391
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 nIleLeuGluGlu-----ThrHisGlnLysIleIleGlyAlaAspGlyGlyIleGlyTh 231
QY 392 TCGGGGTACGCCCCCGCAGGGCTCCTGTGGCAGCGGG-----GCCCCAGCCCAAGCA 445
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
231 rLeuLysArgProLeuAsp---SerGluIleLeuAspGlyAspLeuIleProThrLysIy 250
QY 446 GCAGCAA-----GTGGCAATCCCTCC 466
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
250 sSerSerGluValGlyAspLeuAsnMetGlyAspSerAspLysIleThrAspIle----- 268
QY 467 TCGGCTCTGGTGGCCACCCAGTATGTGGTGCCATTATTCGCAAGAGGGGGGCCACCAT 526
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
269 ----TyrProValProGluLysValGlyLeuValIleGlyLysGlyGlySerGluI1 287
QY 527 CGCAACATCACAAAACAGACCCAGTCCAAAGATAGCGTGCATAGAGAGAGACGACG 586
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
287 eArgLeuIleGlnGlnThrSerGlyCysArgValGlnMetAspProAspHisGlnSerVa 307
QY 587 TGCAGCTGAAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTCTCGCTTGTA 646
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
307 lAsnGlyPheArgAsnCysThrIleGluGlyProProAspGlnValAlaValAlaArgG1 327
```


C;Genetics:
A;Gene: F19C24.19
A;Map position: 1

Alignment Scores:

Pred. No.: 1,71e-10 Length: 621
Score: 265.00 Matches: 117
Percent Similarity: 38.86% Conservative: 74
Best Local Similarity: 23.68% Mismatches: 184
Query Match: 7.50% Indels: 119
DB: 2 Gaps: 19

US-09-270-437D-7 (1-1946) x D96554 (1-621)

```

Qy 459 ATCCCCCTCGCTCTGCTGTCGCCACCCAGTATGTGGTGCATATTGGCAAGAGGGG 518
Db 19 ValHisPheArgLeuLeuGlnCysProAlaThrArgThrGlyAlaIleGlyLysGlyGly 38
Qy 519 GCCACCATCCGACATCACAAAACAGACCCAGTCCAGATA----- 560
Db 39 SerValIleArgHisLeuGlnSerValThrGlySerLysIleArgValIleAspAspIle 58
Qy 560 ----- 560
Db 59 ProValProSerGluGluArgValValLeuIleIleAlaProSerGlyLysLysAsp 78
Qy 561 -----GACGTGCATAGAGAGAGAGACGCGAGTGCAGCTGAA----- 596
Db 79 GluSerAsnValCysAspSerGluAsnProGlySerGluGluProLysGlnGluLysGly 98
Qy 597 -----AAAGCCATCAGTGTGCATCTCCACCCCTGAGGCTCTCTCGCTTGTAA 647
Db 99 SerGluCysAlaGlyThrSerGlyLysAspAspGluGluAlaProSerSerAlaGlnMet 118
Qy 648 ATGATCTGTGAGATTATGCATAAA-----GAGGCTAAGGACACCAAAACGGCTGAC--- 698
Db 119 AlaLeuLeuArgValPheGluArgIleValPheGlyAspAspAlaAlaThrValAspGly 138
Qy 699 -----GAGTTCCCTGAAGATCTCTGGGCCCATTAATACTTT 734
Db 139 AspGluLeuAspLysGlyGluSerGluGlyLeuCysArgMetIleValArgGlyAsnGln 158
Qy 735 GTAGGGGTCTCATTCGCAAGGAGGACGACGACCTGAAGAGGTAGAGCAATACCGAG 794
Db 159 ValAspTyrLeuMetSerLysGlyGlyLysMetIleGlnLysIleArgGluAspSerGly 178
Qy 795 ACAAAATACCATCTCTCGTTGCCA-----GACCTT 827
Db 179 AlaIleValArgIleSerThrAspGlnIleProProCysAlaPheProGlyAspVal 198
Qy 828 ACCCTTTACACCTGAGAGGACCATCATCTGTGAAGGGGCC-----ATCGAGAAT 878
Db 199 ValIleGlnMetAsnGlyLysPheSerSerValLysLysAlaLeuLeuValThrAsn 218
Qy 879 TGT-----TGACGGGCGGACGAGAAATA 902
Db 219 CysLeuGlnGluSerGlyAlaProProThrTrpAspGluCysProPheProGln----- 236
Qy 903 ATGAAGAAAGTTCGGGAGGCTATGAGATGATGCTGCATGCTCCTCCTCCTCCTGATC 962
Db 237 -----ProGlyTyrProProGluTyrHisSerMetGluTyrHis----- 249
Qy 963 CTGGGCTGAACCTGGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG 1022
Db 250 -----ProGlnTrpAspHis-----ProPro 256
Qy 1023 CTTCCAGCAGCGTTACTGGGGTGTCTCCTATAGTCCCTTTATGCGAGGCTCCGAGCAG 1082
Db 257 ProAsnProMetProGluAspValGlyProPheAsnArgProVal-----ValGluGluGlu 275
Qy 1083 GAGATGTGTAGGTGTATTATCCCGCGCAGGCGAGTGGCGCCATCATCGGCAAGAGGGG 1142
Db 276 ValAlaPheArgLeuLeuCysProAlaAspLysValGlySerLeuIleGlyLysGlyGly 295

```

```

Qy 1143 CAGCACATCAAACAGCTCTCCCGGTTTCCAGCGGCTCCATCAAGATTGCACCCGAA 1202
Db 296 AlaValValArgAlaLeuGlnAsnGluSerGlyAlaSerIleLysValSerAspPro--- 314
Qy 1203 ACACCTGACTCCAAAGTTTCGTATGTTATCATCTACTGCGCCGAGAGCCCAATTCAG 1262
Db 315 ThrHisAspSerGluGluArgIleValIleSerAlaArgGluAsnLeuGluArgArg 334
Qy 1263 -----GCTCAGGGA-----AGAACTCTATGCAAACTCAAGGAGGAGAACTTC 1304
Db 335 HisSerLeuAlaGlnAspGlyValMetArgValHisAsnArgIleValGluIleGlyPhe 354
Qy 1305 TTTGGTCCCAAGAGGAGAGTGAAGCTGGAGAGCCACATACATGCTGCGCCAGTACGAGCT 1364
Db 355 -----GluProSerAlaAlaValAlaAlaArgLeuLeuValHisSerProTyrIle 371
Qy 1365 GCGCGGTCTATGTCGAAGGTGGAAACCGTGAACGAGTTGCAGAAATTTCCGCGCAGCT 1424
Db 372 GlyArgLeuLeuGlyLysGlyGlyHisLeuIleSerGluMetArgArgAlaThrGlyAla 391
Qy 1425 GAGGTG---GTAGTACCAAGAGACGACACCCCTGATGAGAACGACCCAG-----GTCA 1475
Db 392 SerIleArgValPheAlaLysAspGlnAlaThrLysTyrGluSerGlnHisAspGluIle 411
Qy 1476 GTGAAATCATCGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATC 1535
Db 412 ValGlnValIleGlyAsnLeu-----LysThrValGlnAspAlaLeuPheGlnIle 428
Qy 1536 CTGCGCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTACACGAGCCAGCGCAGGAGG 1595
Db 429 LeuCysArgLeuArgGluAlaMetPheProGly-ArgLeuProPheGlnGlyMetGlyGly 448
Qy 1596 AAGTGACGACGCCCCCTC-----CCTGCTCCCTTNGAGTCCAGGACA 1634
Db 448 YProProProPheMetGlyProTyrProGluProProProPheGlyProArgGly 468
Qy 1635 ACACGGGCGCAATCGAGAGTGTCTCTCCCGGCGAGC 1674
Db 468 nTyrProAlaSerProAspArgTyrHisSerProValGly 481

```

RESULT 3

T41600
probable pre-mRNA splicing factor - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T41600
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Bothe, G.; Pohl, T.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z22003
A;Accession: T41600
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-398 <LYN>
A;Cross-references: EMBL:AL031825; PIDN:CAA21234.1; GSPDB:GN00068; SPDB:SPCC757.09C
A;Experimental source: strain 972h-; cosmid c757
C;Genetics:
A;Gene: SPDB:SPCC757.09C
A;Map position: 3

Alignment Scores:
Pred. No.: 5.89e-10 Length: 398
Score: 256.50 Matches: 93
Percent Similarity: 44.74% Conservative: 60
Best Local Similarity: 27.19% Mismatches: 130
Query Match: 7.26% Indels: 59
DB: 2 Gaps: 14

US-09-270-437D-7 (1-1946) x T41600 (1-398)

```

Qy 399 CAGCCCGGC-----CAGGGTCCACCTGTGTGCGAGGGGGCCCA----- 437
Db 68 GlnProGluProThrSerGlnValProIleSerAlaLysProPrometAspAla 87

```

QY 371 GCGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACTGTGGCGAGCGG 430
Db 297 -----IleProGlnAsnLeuMe 302
QY 431 GSCCCGAGCAGCAGCAGCAAGTGCACATCCCGCTCGGCTCGTGGTCCCGCACCAGTA 490
Db 302 tMetPro---ArgCysMetLeuLysAspTrpProIleArgCysValValGluGlyLys 321
QY 491 TGTGGGTGCATTATTGGCAAGAGGGGCCACCATCCGCAACATCACAAAAACAGACCCA 550
Db 321 rHisAlaValIleIleGlyProAsnGlySerThrIleLysAspIleAlaSerSerThrAr 341
QY 551 GTCCAGATAGAC-----GTGCATAGGAGGAGAACGCA-----GGTGCAGCTCA 595
Db 341 gCysArgValAspPheValAsnLeuSerLysGluArgThrValLeuGlyAsnAsnAs 361
QY 596 AAAGCCATCAGTGTGCATCTCCACCGCTCAGGCGTCTCTCCGCTGTGTAAGATGATCTT 655
Db 361 pArgIleuThrValHisGlyValAlaGluGlnAlaThrLysAlaValAlaArgIleLe 381
QY 656 GGAGATTATGCATAAAGAGCT-----AAGGACACCAAAACGGCTGACGAGTCCCT 709
Db 381 uAspValIleGlnSerGluAlaValLysAspValAsnValGlyAlaAspThrValLe 401
QY 710 GAAGATCTCGGCCATAATACTTTGTAGCGGTCTCATTTGGCAGGAGGAGCGAACCT 769
Db 401 uArgMetArgAlaHisAsnGlnLeuCysGlyArgLeuIleGlyLysAlaGlySerSerI 421
QY 770 GAAGAGGTGAGCAGCAGATACCGAGACAAAATCACCATCTCC-----812
Db 421 eLysGluIleMetGlnLysThrGlyThrAsnIleThrValThrLysTyrIleGluProPr 441
QY 813 ----TCGTTGCAAGACCTTACC-----CTTTCAACCCCTGAGAGCAGCATCAC 856
Db 441 oGlyGlyIleSerGlyLeuThrAlaAsnGluLeuLeuGlyLeuMetGluArgThrIleMe 461
QY 857 TGTGAAGGG---GCCATCAGAAATTTGTCAGGCGCGCAGCAGGAAATATGAAGAACT 913
Db 461 tValArgGlyProSerIleGluAlaValValGlnAlaGluAlaLeuIleSerAlaLysLe 481
QY 914 TCGGAGGCGCTATGAGAATCATGTGCTGCATGAGCTCTCACCTG-----ATCC 964
Db 481 uLysLysCysTyrGluSerAspSerGlnLeuArgAlaGlnSerMetGlnCysPrometrPr 501
QY 965 TGGCTGAACCTGGCTGTGTAGTCTTTCCAGCTTCATCCAGCGCAGTCCCG-----1019
Db 501 oProMetMetProProIleLeuProGlyAlaSerSerAlaValSerAlaPr 521
QY 1020 -----CGCCCTCCGAGCAGCGCTTACTGGGGTGTCTCCCTAT-----1055
Db 521 oHisPheIleProThrProValGlyValMetGlnValGlnHisPheAlaSerSerGlnHi 541
QY 1056 -----AGTCCTTTATGCAG-----1070
Db 541 sLeuValHisGlnAsnAlaAsnAsnSerPheLeuGlnProGlyValLeuGlnIleGlnPr 561
QY 1071 ----GCTCCGAGCAGAGATGTGAGGTGTTATCCCGCCCGCAGCAGTGGCGCCAT 1126
Db 561 oGlyThrThrAsnLeuArgGlnValArgMetTrpValProAspSerMetIleGlyAlaLe 581
QY 1127 CATCGCAAGAAGGGCAGCAGATCAACAGCTCTCCCGTTTGGCCAGCGCCTCCATCAA 1186
Db 581 uIleGlyAlaLysGlyLysAsnIleLysMetIleArgAspThrGlyAlaSerVally 601
QY 1187 GATTGCACCCGGAACACCTGACTCCAAA-----1217
Db 601 sIleGluAlaProGluGluLysThrGlnArgGluAlaGluGluAlaGluLysLysArgly 621
QY 1217 -----1217
Db 621 sLeuAspGluThrAspSerGlyCysGluGlyValAlaSerGlyAspHisProGlnGluPh 641
QY 1217 -----1217

Db 641 eLeuGluAspAsnAlaThrIleAsnSerSerAspAlaIleGluGluLysProLysProVa 661
QY 1218 ----GTTCTGTATGGTTATCATCTGACCGCCGACAGGCCCAATTCAGGCTCAGGGAAG 1273
Db 661 lSerGluArgMetValThrIleAsnGlyAspAspLeuGlnLeuLysAlaGlnSerTy 681
QY 1274 AATCTATGGCAACTCAAGGAGGAGAACTTTCTTGGTCCC-----1313
Db 681 rValPheSerLysIleAlaGluThrSerSerSerLeuProSerSerGlyMetAspGlyAs 701
QY 1314 ----AAGGAGGAGGAAGTGGAGCACCACATCGTGTGCCAGCATCAGCAGCTGGCGG 1369
Db 701 pArgSerHisMetLeuArgIleArgThrGluValSerValProThrArgIleIleGlyAr 721
QY 1370 GGTCAATTGGCAAGGTGAAACCGTGAACAGTTGCAGAAATTTGACGGCAGCTGAGGT 1429
Db 721 gIleIleGlyLysGlyGlyGlnAsnValArgGluLeuGlnArgIleThrGlyAlaValVa 741
QY 1430 GGTAGTACCAAGACACGACCCCTGAT-----GA 1459
Db 741 lNysIleProGluGluGluArgAsnGlyGlyGluValTyrArgHisAspAspGlyLeuGl 761
QY 1460 GAACCCAGCAGGTCACTCGTGAATAATCATCGGCATTTCTATGCCAGTCCAGATGGCTCAAG 1519
Db 761 uGluAspMetThrMetIleArgThrIleGlyAsnMetTyrSerThrHisAsnValGlnPh 781
QY 1520 GAAGATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATCAGAGGGACAGAGTAACA 1579
Db 781 eArgLeuAlaHisLeu-----ValAsnGluTyrTyrArgSerGlyHisArgAs 798
QY 1580 GCCCAGGACGAGGAGGAGTACACGCCCTCCCTGTCCCTINGACTCAGGACACAC 1639
Db 798 nLysSerSerAspTyrLys-----804
QY 1640 GGGCGAATAATCAGAGTGTGCTCTCCCGCGCAGGCTCAGAAATGAGTGGAAATCCGGGAC 1699
Db 805 -----GlyGlyArgProHisSerAlaProSerSerGlyGl 816
QY 1700 ACNTGGCGCGGCTGTAGATCATGTTGGCTTCCCTACTGATTGAGAAGATGTTCCAGTGAG 1759
Db 816 nGluLysAspGlySerAla-----LeuGluLysMetAspGlnLeuGl 830
QY 1760 AACCTGTATCTTCAGCCCCAACCCCAATTTGCCCAACACTGNTGCCCTCG 1819
Db 830 yThrIle-----AlaProLysSerAsnSerAsnArgAlaSerProLysSe 845
QY 1820 GGTGTCA 1826
Db 845 rValSer 847
RESULT 2
D96554
hypothetical protein F19C24.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96554
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luross, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96554
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-621 <STO>
A:Cross-references: GB:AE005173; NID:g11094762; PIDN:AAG29695.1; GSPDB:GNC0141

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:55:36 ; Search time 29.4003 Seconds
(without alignments)
12733.810 Million cell updates/sec

Title: US-09-270-437D-7

Perfect score: 3531

Sequence: 1 gctgtgcgagggcgtggg.....attctctcaggttttaaaa 1946

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0			0.5
Ygapop 5.0			7.0
Deiop 6.0			Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO pool p/US09270437/runat 23072004.095740.27439/app query.fasta_1.5582
-DB=PIR_78 -QFMT=fastan -SUFFIX=xpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORES=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HBAPEXT=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1 1 134 @runat 23072004.095740.27439 -NCPUS=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	458.5	13.0	854	2 T23837	hypothetical prote
2	265	7.5	621	2 D96554	hypothetical prote
3	256.5	7.3	398	2 T41600	probable pre-mRNA
4	255	7.2	768	2 T27855	hypothetical prote
5	251	7.1	479	2 C86275	hypothetical prote
6	244	6.9	644	2 A53184	myc far upstream e
7	240.5	6.8	680	2 T25832	hypothetical prote
8	234.5	6.6	568	2 T49962	hypothetical prote
9	232.5	6.6	632	2 T02627	hypothetical prote
10	232	6.6	589	2 T19216	hypothetical prote
11	232	6.6	611	2 T19217	hypothetical prote
12	221.5	6.3	510	2 T38489	onconeural ventral
13	218.5	6.2	356	2 S58529	alpha-complex prot
14	218	6.2	1268	2 A44125	high density lipop

15	217.5	6.2	396	2 S41224	hnRNP protein - Af
16	214	6.1	313	2 T48439	probable RNA-bind
17	209	5.9	397	2 T30168	hypothetical prote
18	205.5	5.8	846	2 T04533	hypothetical prote
19	202.5	5.7	1270	2 S23464	vigilin - chicken
20	200.5	5.7	1279	2 T41389	rna binding protei
21	199	5.6	365	2 S42471	hnRNP protein E2 -
22	197.5	5.6	463	2 S41495	dc stretch-binding
23	197	5.6	464	2 S43363	transformation upr
24	194	5.5	362	2 S78515	single-stranded nu
25	193	5.5	464	2 A54143	kappa-B motif-bind
26	191	5.4	413	2 S46109	hnRNP complex prot
27	190.5	5.4	1220	2 T19117	hypothetical prote
28	190	5.4	1198	2 B88279	protein C08H9.2 li
29	182.5	5.2	1222	2 S56030	SCP160 protein - y
30	178	5.0	381	2 S45766	hypothetical prote
31	165	4.7	806	2 T13690	hypothetical prote
32	164	4.6	784	2 J00317	hypothetical 82K p
33	162.5	4.6	848	2 S02262	glutenin high mole
34	161	4.6	891	2 G84693	probable proline-r
35	159	4.5	389	2 C96460	FI4M2.18 protein -
36	156.5	4.4	648	2 S04832	glutenin high mole
37	156.5	4.4	1283	2 T49692	related to SCP160
38	155.5	4.4	228	2 S42933	hypothetical prote
39	155.5	4.4	255	2 J00320	hypothetical 24.7K
40	153	4.3	1460	1 EDBE1F	immediate-early pr
41	153	4.3	1872	2 S36152	MHC class III hist
42	151.5	4.3	495	2 T04255	hypothetical prote
43	151.5	4.3	2715	2 T13049	eyelid - fruit fly
44	151	4.3	838	1 EEWTHW	glutenin, high mol
45	150	4.2	2142	2 B35098	MHC class III hist

ALIGNMENTS

RESULT 1

T23837 hypothetical protein M88.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T23837

R:Salston, J.

submitted to the EMBL Data Library, June 1994

A:Reference number: Z19806

A:Accession: T23837

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-854 <WIL>

A:Cross-references: EMBL:Z34802; PIDN:CAA84338.1; GSPDB:GN00021; CESP:M88.5

A:Experimental source: clone M88

C:Genetics:

A:Gene: CESP:M88.5

A:Map position: 3

A:Introns: 38/3; 96/1; 133/3; 255/3; 307/3; 395/2; 436/2; 492/3; 657/3; 759/1; 820/1

Alignment Scores:

Pred. No.:	5.99e-24	Length:	854
Score:	458.50	Matches:	160
Percent Similarity:	39.04%	Conservative:	91
Best Local Similarity:	24.88%	Mismatches:	212
Query Match:	12.98%	Indels:	180
DB:	2	Gaps:	21

US-09-270-437D-7 (1-1946) x T23837 (1-854)

QY	254	CAACCGGAGCAGACGAGCAAGCCATCATGAAGCTGAATGCCAC---CAGTTGGAGAA	310
DB	267	GlnAlaGlnHlnHisGlnGlnSerHisGlnSerGln-AsnHisAsnGlnHisArgAs	286
QY	311	CCATGCCCTGAAGGCTCTCTACATCCCGCATGACAGATAGCACAGGGACCTGGAATGG	370
DB	286	nHis-----AsnGlnSerHisSerGlyProHisHis--	296


```
QY 753 AAGGAAGACCGAACCCTGAGAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATTCTCC 812
Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTGTGAAGACCTTACCCTTTCAACCCCTGAGAGGCACCTGCTGGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTGTTTCAGCGGCGAGCAGCAATAATGAAAGATTTCGGGAGGCCCTATGAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrgluAsn 353
QY 933 GATGTGGTGCATGAGC-----TCTCACTGATCCCTGGCTGGAACCTGGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCNCCAGCAGCGCTT 1037
Db 374 GlyLeuPheProThrSerSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCCTATAGCTCTCTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTG 1097
Db 394 Thr-----ProProTyrgluGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCCGCAGGAGTGGGCCCATCATGCGCAAGAGAGGGGCGACACATCAAAAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGTGTTCAGCGCTCCATCAAGATTGCACCCAGCCGCAACACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAGGAGAAATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgile 470
QY 1278 TATGGCAAACTCAAGGAGAGAACTTTTGTGTCCTCAAGAGGAAAGTGAAGCTGGAGACC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCCAGCATCAGCAGCTGGCGGGTCATTGGCAAGGTGGAAACGCTG 1397
Db 491 HisIleArgValProSerPheAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
QY 1398 AACGAGTTCGCAATTTGACGGCAGCTGAGTGTGTGTTACCAAGACAGCAGACCCCTCAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACCGACCGTTCATCGTGAATAATCATCGACATTTCTATGCGCAGTCCAGATGCTCAA 1517
Db 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyrglyAlaCysGlnValAlaGln 550
QY 1518 CGGAGATCCGAGACATCTGCGCCCGAGGTAAAGCAG---CAGCATCAGAAAGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGGCGCCAGCGCAGGAGGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
```

RESULT 15

```
US-10-313-986-348
; Sequence 348, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
```

```
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-313-986-348

Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 15 Gaps: 9

US-09-270-437D-7 (1-1946) x US-10-313-986-348 (1-579)
```

```
QY 102 CGGTTAGGAGCCGGAATAATTCAAATCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTyrGluVal 96
QY 162 CTGACAGCCCTGCTGGCTCAGTATGGTACGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGCGGAGTGGTGAATGTCACTTATTCACCCGCGGAGCAGACCCAGGCAAGCCATC 281
Db 117 SerGluThrAlaValAlaValAsnValThrTySerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACCAGTGGAGAACCATGCTCCCTGAAAGTCTCTTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCCCGCGAGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGTTCAGCCCGCCGCGGCTCAGCTGTGGCAGCGGGCCCGCCAGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGACATPCCCTTCGCTCCTGTGCTGCCACCCAGTATGTGGTGCCATTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACCGAGGTGAGCTGAAAGCCATCAGTGTGCTCCACCCCTCAGGCGCTGC 632
Db 234 LysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTGATAGATGATCTTGAGATATTATGCAATAAGAGCTTAAGACACCAAGCG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCCTGAAAGATCTGGCCCATATAACTTTGTAGGGCGTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGCAGCAACCTCAAGAGGTAGCAAGATACCGACACAAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTACCCCTTTACACCCCTGAGAGGACCATCCTGTGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAsnVal 333
```

```
QY 693 GCTGACGAGGTTCCCTGAGATCCCTGGCCCAATAAATCTTTGTAGGGCTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuLeuGly 293
QY 753 AAGGAGGCGGAACTGAGAGGTAGACAGATACCCAGACAAATACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGluAspThrAspThrLysIleThrIleSer 313
QY 813 TCCTGCAACACCTTACCTTTTCAACCTCTGAGAGGACCATCTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTGTCAGGGCCGAGCGAGAAATATGAGAAATTCGGAGGCGCTTAGAAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTGCCATGAGC-----TCTCACCTGATCCTGGCTGACCTGGCTGTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTCTCCTTAGCTCTTTATGAGCTCCCGAGGATGGTGCGAGT 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGGCGGCGCATCATCGGAAGGCGGAGCAGCATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln 430
QY 1158 CTCCTCCGGTTGCCAGCGCTCCATCAGATTGCACCCAGCCCGAAACACCTGCTCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCATGACCGCCGAGCGCCCAATTCAGGCTCAGGAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAGGAGGAGAACTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGAC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGCTGTCAGCATCAGCAGCTGGCGGCTCATTCGGCAAAAGGTGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTCAGATTTGCGGCGAGCTGAGTGTGTAGTACCAAGACGACGACCTGTAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACACGACGAGTTCATCGTGAATAATCATCGACATTTCTATGCGCATCGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGGCCAGGTAAAGCAG---CAGCATCAGAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACACGCGCCAGGACGAGGAAG 1598
Db 571 SerGlyProProGlnSerArgLys 579
```

RESULT 14

```
US-10-117-982-480
; Sequence 480, Application US/10117982
; Publication No. US20030139438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
```

```
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 480
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-480
```

```
Alignment Scores:
Pred. No.: 1-52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 14 Gaps: 9
```

US-09-270-437D-7 (1-1946) x US-10-117-982-480 (1-579)

```
QY 102 CGGTTAGAGAGCGGGAATTCAAATCCGAATATTCACCCCGAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTCGACAGCTGCTGCTCAGTATGTCACAGTACAGAACTGTGAGCAAGTGAACACGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGlnGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGTGATGTCCTATTCACCGGGAGCAGACCCAGGCAAGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCGCCACAGTGTGGAGAACCATGCCCTGAGTCTCCATCATCCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTGAGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGACATCCCTCTCGGCTCTGCTGGTGGCCACCCAGTATCTGGTGGTCCATTATTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCCAAGATAGAGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGTGCAGCTGAAAGGCCATCAGTGTGCTCCACCCCTGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCCTGAAAGATCCTGGGCCCAATAAATCTTTAGGCGCTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuLeuGly 293
```

```

Db      254 SerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGlnAsnIleLysPhe 273
Qy      693 GCTGACGAGGTTCCCTGAAAGTCTGCGCCCATATAACTTTGTAGGCGCTCTATTGGC 752
Db      274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
Qy      753 AAGGAAGGACGGAACCTGAAGAGGTAGAGAAGATACCGAGACAAAAATCACCATCTCC 812
Db      294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
Qy      813 TCGTTGCAAGACCTTACCTTTACAACTCAGAGGACCATCACTGTGAAGGGGCCATC 872
Db      314 ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValGlyAsnVal 333
Qy      873 GAGAATTTCTGCGAGGCGGACGAGGAATAATGAAGAAAGTTTCGGGAGCGCTATCAGAA 932
Db      334 GluThrCysAlaLysAlaGluGluIleMethLysLysIleArgGluSerTyraGluAsn 353
Qy      933 GATGTGGTGCATGAGC-----TCTCACTGTATCTCTGGCTGAACTGGCTCTCTGTA 986
Db      354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
Qy      987 GGTCTTTTCCAGCTTCACTCAGCGCAGCTCCGCGC-----CCTCCGACGAGCCTT 1037
Db      374 GlyLeuPheProProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
Qy      1038 ACTGGGGTGTCTCCCTATAGCTCTTTATGACAGGCTCCGAGCAGGAGATGTCGAGGTG 1097
Db      394 Thr-----ProProTyraProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
Qy      1098 TTTATCCCGCCGAGGAGTGGGCGCCATCATTCGGCAAGAGAGGGGCGACATCAAAAG 1157
Db      411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
Qy      1158 CTCTCCCGTTTGGCAGCGCTCCATCAAGATTGACACCCGCAACACCTGACTCCAAA 1217
Db      431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
Qy      1218 GTTCGTATGGTTATCATCACTGGACCGCGCAGAGGCGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db      451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
Qy      1278 TATGCAAACTCAAGAGGAGAACTTTCTTGTCCCAAGGAGGAAGTGAAGCTGAGACC 1337
Db      471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
Qy      1338 CACATACGTGTGCCAGCATCAGCAGCTGCGCGGTTCATTGGCAGAGGTGGAAACGCTG 1397
Db      491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
Qy      1398 AACGAGTTGCAAGATTTGACGCGCAGCTGAGGTGGTGTAGTACCAAGAGACGACGCCCTGAT 1457
Db      511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
Qy      1458 GAGAACGACGAGTCACTCGTGAATAATCATCGACATTTCTATGCCAGTCAGATGCTCAA 1517
Db      531 GluAsnAspGlnValValLysIleThrGlyHisPheTyraLysGlnValAlaGln 550
Qy      1518 CGGAAGATCCGAGACATCTCGCCCGAGGTAAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db      551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
Qy      1572 AGTAACACGAGCGCCGACGAGGGAAG 1598
Db      571 SerGlyProProGlnSerArgArgLys 579

```

RESULT 13

US-10-117-982-449
; Sequence 449, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.

```

; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; SOFTWARE: FASTSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 484
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-449

Alignment Scores:      1.52e-133      Length:      579
Pred. No.:            1875.00      Matches:      380
Score:                84.68%      Conservative:  51
Best Local Similarity: 74.66%      Mismatches:   62
Query Match:          53.10%      Indels:       16
DB:                   14          Gaps:         9

```

US-09-270-437D-7 (1-1946) x US-10-117-982-449 (1-579)

```

Qy      102 CGGTTAGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
Db      77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
Qy      162 CTGACAGCCTGCTGGCTCAGTATGGTACAGTAGAAGACTGTGACCAAGTGNACACCGAG 221
Db      97 LeuAspSerLeuLeuValGlnTrpGlyValGluSerCysGluGlnValAsnThrAsp 116
Qy      222 AGTGAGAGCGGAGTGGTGAATGTCACTATTCCACCGGAGCAGACACGAGCAAGCCATC 281
Db      117 SerGluThrAlaValAlaValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeu 136
Qy      282 ATGAGCTGAATGGCCACCACTGGAGAACCATGCTCCCTCAAGGTCTCTACATCCCGAT 341
Db      137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaIleProasp 156
Qy      342 GAGCAGATAGCA---CAGGACCTGAGAAATGGGCGCCGAGG-----GGCTTGGCTCT 392
Db      157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
Qy      393 CGGGTTCAGCCCGCCGAGGCTCACCTGTGGCAGCGGGCCCGCCAGCAAGCAGCAGCAA 452
Db      177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
Qy      453 GTGACATCCCTTCGGCTCTCTGTGTCGCCACCCAGTATGTGGTGCCATTATGGCAAG 512
Db      194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
Qy      513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGACGTGCATAGG 572
Db      214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrLysGlnSerLysIleAspValHisArg 233
Qy      573 AAGGAGAACGCGAGGTGAGCTGAAAGCCATCACTGTGTGCACTCCACCCCTGAGGCTGC 632
Db      234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
Qy      633 TCCTCCGCTTGTAAAGATGATCTTGAGATTATGCATAAAGAGCTTAAGACACCAAGAG 692
Db      254 SerAlaAlaCysLysSerIleLeuGluIleMethHisLysGluAlaGlnAspIleLysPhe 273

```



```
Db      234  LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY      633  TCCTCCGCTGTAGATGATCTTGGAGATTATGATTAAGAGGCTAAGGACACCAAAACG 692
Db      254  SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY      693  GCTGACGAGGTCCCTCGAGATCCTCGCCCATATTAACCTTGAGGGCGCTCATGGC 752
Db      274  ThrGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY      753  AAGGAAGCGAGAACCTCAAGAGGTAGAGCAAGATACCGAGACAAATAACCATCTCC 812
Db      294  LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY      813  TCGTTGCAAGACCTTACACCTTGAGAGGACCATCACTGTGAAGGGGGCCATC 872
Db      314  ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY      873  GAGAATTGTTGCGAGCGCAGCAGGAATAATGAAGAAAGTTCGGGAGGCGCTATGAGAAT 932
Db      334  GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgLysSerTyrGluAsn 353
QY      933  GATGTGGCTGCCATGAGC-----TCTCAGCTGATCCCTGCTGAACTGGCTGTGTA 986
Db      354  AspIleAlaSerMetAsnLeuGlnAlaHisIleLeuProGlyLeuAsnLeuAlaLeu 373
QY      987  GGTCTTTTCCAGCTTCATCCAGCGCAGTCCGCGC-----CCTCCAGCAGCGTT 1037
Db      374  GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY      1038  ACTGGGGCTGCTCCTATAGCTCTTATGTCAGGCTCCCGAGCAGGAGATGTCGAGTG 1097
Db      394  Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY      1098  TTATATCCCGCCAGCGAGTGGCGCCATCATCGCAAGAGAGGGCAGCACATCAACAG 1157
Db      411  PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY      1158  CTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCCCGAAGAACCTGATCTCAAA 1217
Db      431  LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY      1218  GTTCGTATGTTATCATCACTGACCGCCAGAGCCCAATTCAAGGCTCAGGAGGAATC 1277
Db      451  ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY      1278  TATGGCAAACTCAGGAGGAGAACTTCTTTGGTCCCAAGAGGAGTGAAGCTGGAGACC 1337
Db      471  TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY      1338  CACATAGCTGTCAGCATCAGCAGCTGCGCGGTGATTTGGCAAGGTGAAAAACGGTG 1397
Db      491  HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY      1398  AACGAGTTGAGAAATTTGACGCGAGCTGAGTGTAGTACCAAGAGACCAAGCCCTGAT 1457
Db      511  AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY      1458  GAGAAGCAGCAGTFCATCGTGAATAATCATCCGACATTTCTATCCAGTCAAGTGGCTCAA 1517
Db      531  GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY      1518  CGGAGATCCGAGACATCTCGCCAGGTAAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db      551  ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnValAlaLeuGln 570
QY      1572  AGTAACCCAGGCGCCAGCAGGAGAG 1598
Db      571  SerGlyProProGlnSerArgArgLys 579
```

RESULT 12

US-10-117-982-446

; Sequence 446, Application US/10117982

```
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: FOY, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-446

Alignment Scores:
Pred. No.:      1,52e-133      Length:      579
Score:          1875.00      Matches:    380
Percent Similarity: 84.68%      Conservative: 51
Best Local Similarity: 74.66%      Mismatches: 62
Query Match:      53.10%      Indels:     16
DB:              14          Gaps:        9

US-09-270-437D-7 (1-1946) x US-10-117-982-446 (1-579)
```

```
QY      102  CGGGTTAGAGCGCGAAATTCGAATCCGNAATATTCACCCCGAGCTCCGATGGGAAGTA 161
Db      77  ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTyrGluVal 96
QY      162  CTGACAGACCTGCTGCTCAGTATGTACAGTAGAGAACCTGTGAGCAAGTGAACACCGAG 221
Db      97  LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY      222  AGTGACAGCGGAGTGTGATGTCACCTATTCACACGGGAGCAGACCGAGCGAAGCATC 281
Db      117  SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY      282  ATGAAGCTGAATGGCCACCGAGTTGGAGAACCATGCCCTGAGGCTCTCTACATCCCGAT 341
Db      137  AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY      342  GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db      157  GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY      393  CGGGGTGACCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAAA 452
Db      177  ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY      453  GTGGACATCCCTCTCGGCTCCTGCTGTCGCCACCCAGTATGTGGGTGCCATTTATGGCAAG 512
Db      194  CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY      513  GAGGGGGCCACCATCCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db      214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY      573  AAGGAAGCAGCGGTGACGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGC 632
Db      234  LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY      633  TCCTCCGCTTGTAAAGTATGATCTTGGAGATTATGCATAAAGAGCGCTAAGACACCAACG 692
```

RESULT 11
 US-10-117-982-348
 : Sequence 348, Application US/10117982
 : Publication No. US2003013438A1
 : GENERAL INFORMATION:
 : APPLICANT: FOV, Teresa M.
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Vedvick, Thomas S.
 : APPLICANT: Carter, Darrick
 : APPLICANT: Watanabe, Yoshihiro
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Maricle, Barbara
 : APPLICANT: Spies, Gregory A.
 : APPLICANT: Fan, Liqun
 : APPLICANT: Wang, Tongtong
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 : FILE REFERENCE: 210121.455C18
 : CURRENT APPLICATION NUMBER: US/10/117,982
 : CURRENT FILING DATE: 2002-04-05
 : NUMBER OF SEQ ID NOS: 484
 : SOFTWARE: Fast-Seq for Windows Version 4.0
 : SEQ ID NO 348
 : LENGTH: 579
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-10-117-982-348
 Alignment Scores:
 Pred. No.: 1,52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 14 Gaps: 9
 US-09-270-437D-7 (1-1946) x US-10-117-982-348 (1-579)
 QY 102 CGGGTTAGGAGCGGAAATTCACCGAAATATTCACCCAGCTCCGATGGAAAGTA 161
 Db 77 ArgGlnArgIleArgIysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCCCTGCTGGCTCAGTATGTGTACAGTACAGAACTGTGAGCAAGTGAACACCCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTCAGAGCGGAGTGTGTGAATGTGCACCTATTCCACCGGAGCAGACCGCAGCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCACTGTGGAGAACCATGCCCTGAAGTCTCTATCATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGGCCGAGGG-----GGCTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCCAGGCTACCTGTGGCAGCGGGGGCCCGACGAGCAGCAGCA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTTCGGCTCTGGTGCACCCAGTATGGTGGCTATTTGGTGGCAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleGlyLys 213
 QY 513 GAGGGGCCCATCCCGCAATCATCAAAACAGACCCAGTCCCAAGATAGAGTCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACGCGAGTGCAGCTGAAAGAACCATCATGTGTGCACTCCACCCCTGAGGGTGC 632
 573 AAGGAGAACGCGAGTGCAGCTGAAAGAACCATCATGTGTGCACTCCACCCCTGAGGGTGC 632
 : Sequence 348, Application US/10117982
 : Publication No. US2003013438A1
 : GENERAL INFORMATION:
 : APPLICANT: FOV, Teresa M.
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Vedvick, Thomas S.
 : APPLICANT: Carter, Darrick
 : APPLICANT: Watanabe, Yoshihiro
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Maricle, Barbara
 : APPLICANT: Spies, Gregory A.
 : APPLICANT: Fan, Liqun
 : APPLICANT: Wang, Tongtong
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 : FILE REFERENCE: 210121.455C18
 : CURRENT APPLICATION NUMBER: US/10/117,982
 : CURRENT FILING DATE: 2002-04-05
 : NUMBER OF SEQ ID NOS: 484
 : SOFTWARE: Fast-Seq for Windows Version 4.0
 : SEQ ID NO 348
 : LENGTH: 579
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-10-117-982-348
 Alignment Scores:
 Pred. No.: 1,52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 14 Gaps: 9
 US-09-270-437D-7 (1-1946) x US-10-117-982-348 (1-579)
 QY 102 CGGGTTAGGAGCGGAAATTCACCGAAATATTCACCCAGCTCCGATGGAAAGTA 161
 Db 77 ArgGlnArgIleArgIysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCCCTGCTGGCTCAGTATGTGTACAGTACAGAACTGTGAGCAAGTGAACACCCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTCAGAGCGGAGTGTGTGAATGTGCACCTATTCCACCGGAGCAGACCGCAGCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCACTGTGGAGAACCATGCCCTGAAGTCTCTATCATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGGCCGAGGG-----GGCTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCCAGGCTACCTGTGGCAGCGGGGGCCCGACGAGCAGCAGCA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTTCGGCTCTGGTGCACCCAGTATGGTGGCTATTTGGTGGCAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleGlyLys 213
 QY 513 GAGGGGCCCATCCCGCAATCATCAAAACAGACCCAGTCCCAAGATAGAGTCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACGCGAGTGCAGCTGAAAGAACCATCATGTGTGCACTCCACCCCTGAGGGTGC 632
 573 AAGGAGAACGCGAGTGCAGCTGAAAGAACCATCATGTGTGCACTCCACCCCTGAGGGTGC 632
 : Sequence 348, Application US/10117982
 : Publication No. US2003013438A1
 : GENERAL INFORMATION:
 : APPLICANT: FOV, Teresa M.
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Vedvick, Thomas S.
 : APPLICANT: Carter, Darrick
 : APPLICANT: Watanabe, Yoshihiro
 : APPLICANT: Henderson, Robert A.
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Maricle, Barbara
 : APPLICANT: Spies, Gregory A.
 : APPLICANT: Fan, Liqun
 : APPLICANT: Wang, Tongtong
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 : FILE REFERENCE: 210121.455C18
 : CURRENT APPLICATION NUMBER: US/10/117,982
 : CURRENT FILING DATE: 2002-04-05
 : NUMBER OF SEQ ID NOS: 484
 : SOFTWARE: Fast-Seq for Windows Version 4.0
 : SEQ ID NO 348
 : LENGTH: 579
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 : US-10-117-982-348
 Alignment Scores:
 Pred. No.: 1,52e-133 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 14 Gaps: 9
 US-09-270-437D-7 (1-1946) x US-10-117-982-348 (1-579)
 QY 102 CGGGTTAGGAGCGGAAATTCACCGAAATATTCACCCAGCTCCGATGGAAAGTA 161
 Db 77 ArgGlnArgIleArgIysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGACAGCCCTGCTGGCTCAGTATGTGTACAGTACAGAACTGTGAGCAAGTGAACACCCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTCAGAGCGGAGTGTGTGAATGTGCACCTATTCCACCGGAGCAGACCGCAGCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCACTGTGGAGAACCATGCCCTGAAGTCTCTATCATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGGCCGAGGG-----GGCTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCCAGGCTACCTGTGGCAGCGGGGGCCCGACGAGCAGCAGCA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTTCGGCTCTGGTGCACCCAGTATGGTGGCTATTTGGTGGCAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleGlyLys 213
 QY 513 GAGGGGCCCATCCCGCAATCATCAAAACAGACCCAGTCCCAAGATAGAGTCATAGG 572
 Db

```
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTTGTAAAGATGATCTTCGAGATTATGATAAAGAGGCTAAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACAGGTTCCCTCCAGATCTGCGCCCATATAACTTTGTAGGCGTCTCATGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGGACGGAACCTGGAAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrLysIleThrIleSer 313
QY 813 TCGTGTGAAGACCTTACCTTTACAACCTGAGAGGACCATCACTGTGAAGGGGCCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGRAATTGTTCCAGGCGGACGAGCAAGGAATATGAGAAAGTTCGGGAGGCTATGAGAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGTGCATGAGC-----TCTCACCTGATCCTGCGCTGAACCTGGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCATGCTCCGCG-----CTCCACGAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProSerSerAlaMet 393
QY 1038 ACTGGGCTGCTCCTATAGTCTCTTATGAGCTCCCGAGGAGGATGTCGAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCAGGAGTGGCGCCCATCATCCGCAAGAGGGCGAGCATCAAAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCGTTGTCAGCGCTCCATCAAGATTGCACACCCGAAACCTGTGATCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTGTGTGTATCATCACTGGACCGCCAGAGCCCAATTCAGGCTCAGGAGGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGAGGAGAACTCTTTGGTCCCAAGAGGAAGTGAAGCTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGTGTCCAGCATCAGAGCTGGCGGGTCTATGGCAAGGTGGAANAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGATTTCAGCGAGCTGAGTGTAGTACCAAGAGACCAAGCCCTCAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACACGAGGTTCATCTGTGAATAATCATCGACATTTCTATGCGAGTCAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAGATCCGAGACATCTCGGCCCGAGTTAAGAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACACGAGCGCCAGGACGAGGAAG 1598
Db 571 SerGlyProProGlnSerArgLys 579
```

RESULT 10

US-10-007-700-449

; Sequence 449, Application US/10007700

```
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Peng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-007-700-449

Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 12 Gaps: 9

US-09-270-437D-7 (1-1946) X US-10-007-700-449 (1-579)
```

```
QY 102 CGGGTAGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCCTGCTGGCTCAGTATGTAGTACAGAACTGTGAGCAAGTGAACACCCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGTGTAATGCACCTATTCCAACCGGAGCAGACCCAGGCAAGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaAsgGlnAlaLeu 136
QY 282 ATGAGCTGAATGCCACACAGTTGGAGAACCTACCTGAGGCTCTCTACATCCCGCAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgargGlyLeuGlyGln 176
QY 393 CGGGGTGAGCCCGCAGGCTCACCTGTGGCGAGCGGGGCCCGCCAGCAGGAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGAGATCCCTCGCTCGCTCTGCTGGTGGCCACCCAGTATGTGGTGGCTATTTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
```

```

Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCTCCAGATCTCGCCCATATAACTTTGTAGGGCGTCTCATTCGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGACCGAAGCTGAAGAGGTAGACGAAGATACCGACAAATAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTCAGACCTTACCTTTACACCTCGAGGACCATCCTGTGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAATTCGTGACGGCCGACGAGGAATAATGAAGAAATTCGGGAGCGCTTATCAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGTGCATGAGC-----TCTCACTGTATCCTCGCTGAACTGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCAGCCAGTCCCGCG-----CCTCCGAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCCTATAGCTCTTATGAGGCTCCCGAGCAGGATGTCAGGTTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGGAGTGGCGCCATCATTCGCAAGAGAGGGCAGACATCAAAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGTTGCGAGCGCTTCATCAGATTGACACCCGAAACACCTGACCTCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCTATGTTTATCATCATGAGCGCGCAGAGGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGCCAAACTCAAGGAGAGAACTTCTTTGTTCCCAAGAGAGAAAGTGAGTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGCGCAGATCAGCAGTGGCGGCTCATTCGCAAGAGTGGAAAGCGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTCAGCGCAGCTGAGGTGGTAGTACCAAGAGACCCAGCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACCGAGTTCATCGTGAATAATCATCGACATTTCTATCCAGTCCAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGATCTCTGCCCGCAGGTAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAAACACGCGCCGAGCAGGAGGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579

```

RESULT 9

US-10-007-700-446
; Sequence 446, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

```

; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-446

Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 12 Gaps: 9

US-09-270-437D-7 (1-1946) x US-10-007-700-446 (1-579)
QY 102 CGGGTTAGGAGCGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGAGCGGAGTGTGAATGTCACTTATTCACCGGAGCAGACAGGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCGCTGAGAGGTCTCTTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGATGGCGCCGAGGG-----GGCTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGln 176
QY 393 CGGGGTGACGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCGGAACATCAAAACAGACCCAGTCCCAAGATGACATGCTGATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGGTGCGAGCTGAAAAGCCATCAGTGTGCTACCTCCACCCCTGAGGCTGC 632

```

```
QY 693 GCTGACAGAGTTCCCTGAGATCTCTGGCCCAATAAATCTTTGTAGGGCTCTCATTTGGC 752
Db 274 ThrGluGluLeuProLeuLysIleLeuAlaHisAsnAsnProValGlyArgLeuLysGly 293
QY 753 AAGGAGGCGGACCTTGAAGAGGTAGACAGATACCGAGACCAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrIleThrIleSer 313
QY 813 TCCTTGCAGACCTTACCCCTTTTACAACCTCGAGAGCACCATCACTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATGTTGCGGGCGGACGAGCAATAATAGAAAGTTCCGGAGGCTATGAGAAAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTCCATGAGC-----TCTCACCTGATCCCTGGCTGAACCTTGGCTGTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGCTCTTTCCAGCTTCATCCAGCGCATCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTGCTCCCTATAGCTCTTTATGAGGCTCCCGAGCGAGATGGTGCAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCCGGAGGCGGCGCATCATCGGCAAGAGGGGAGCACATCAAAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleGlyLysGlnGlyGlnHisIleLysGln 430
QY 1158 CTCTCCCGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCACTGACCGCCAGAGGCCCAATTCAGGCTCAGGAGAAATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAGGAGGAGAACTCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGCTGTCAGCATCAGCAGCTGGCGGGTCAATTCGGCAAAAGTGAAGAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTTGACGCGAGCTGAGTGTAGTACCAAGAGACGAGCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACGAGGTTCATCGTGAATAATCATCGGCATTTCTATGCACTCAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCTGGCCCGCAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGAGCCCGGAGGAGGAG 1598
Db 571 SerGlyProProGlnSerArgLys 579
```

RESULT 8

```
US-10-007-700-348
; Sequence 348, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
```

```
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Peng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-348
```

```
Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 12 Gaps: 9
```

US-09-270-437D-7 (1-1946) x US-10-007-700-348 (1-579)

```
QY 102 CGGGTTAGAGCCCGGAAATTCAAATCCGAAATATTCACCCCGAGCTCCGATCGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCCTGCTGGCTCAGTATGTGACAGTGTGAGAACCTGTGAGCAAGTGAACCCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGlnGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGCTGGTGAATGTCACCTATTCCAACCGGAGCAGACCCAGGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATCAAGCTGAATGCCACCCAGTTGGAGAACCATGCCCTGAAGGCTCTCCTACATCCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCGCAGGG---GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTGAGCCCGCAGGCGCTCACCTGTGGCAGCGGGGCGCCAGCAGAGAGAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCCCTCGGCTCCTGGTGGCCACCCAGTATGTGGTGGCTTATTTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCGCTGTGAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGACACCAACG 692
```



```
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
Qy 513 GAGGGGCGCCATCCGCAACATCACAAACAGACCCAGTCCAGATAGAGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
Qy 573 AAGGAGAACCGAGTGCAGCTGAAAGAACCCATCATGTGCTCCACCCCTGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
Qy 633 TCCTCCGCTTATAGATCATCTTGGAGTTATGATTAAGAGGCTTAAGGACACCAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
Qy 693 GCTCAGAGGTTCCCTGAAAGTCTGCGCCCATATAACTTTGTAGGGCGTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
Qy 753 AAGGAGGACGGAACCTGAAGAGGTAGAGCAAGATACGAGACAAATAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
Qy 813 TCGTTGCAAGACTTACCTTTACACCTGAGGACCATCATCTGTAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValLysGlyAsnVal 333
Qy 873 GAGAATTGTTCCAGGCGGAGCAGGAATAATGAAGAAAGTTCCGGAGGCTTATCAGAAAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyraGluAsn 353
Qy 933 GATGTGGCTGCATGAGC-----TCTCACTGATCCCTGGCCTGAACCTGGCTCTCTTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
Qy 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
Qy 1038 ACTGGGGTGTCTCCTATAGCTCTTATGAGGTCCTGAGGTCCTGAGGAGATGTCGAGGTG 1097
Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
Qy 1098 TTTATCCCGCCCGCAGCAGTGGCGCCATCATCGCAAGAGAGGCGAGCACATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln 430
Qy 1158 CTCTCCCGGTTGCCAGCGCTCCATCAAGATTGCAACCCGCAACACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
Qy 1218 GTTCGTATGTTTATCATCTGACCGCGCAGAGGCCCAATTCAAGCCTCAGGGAAGAATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
Qy 1278 TATGCCAACTCAAGAGAGAACTCTTTGGTCCCAAGAGAGAGTGAAGCTGAGAC 1337
Db 471 TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
Qy 1338 CACATACGTGTCAGCATCAGCAGCTGCGCGGTTCATTGGCAAGAGTGGAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
Qy 1398 AACGAGTTGCAGAAATTTGACCGCAGCTGAGGTGGTAGTACCAAGAGACACACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
Qy 1458 GAGAACGACAGGTGCATCGTGAATTCATCGACATTTCTATGCCAGTCAAGTGCCTAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyraLysGlnValAlaGln 550
Qy 1518 CGAAGATCCGAGATCCTGCGCCAGGTTAGCAG---CAGCATCAGAGGGA---CAG 1571
```

```
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
Qy 1572 AGTAACAGAGCCCGACGACGAGGAGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
RESULT 5
US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, TongTong
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-348
Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 9 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-897-778-348 (1-579)
Qy 102 CGGTTAGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
Qy 162 CTGGACAGCTGCTGGCTCAGTATGATAGTAGAAGTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAsp 116
Qy 222 AGTGAGACGCGAGTGGTGAATGTCACTTATTCACACCGGGAGCAGACGAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyraSerSerLysAspGlnAlaArgGlnAlaLeu 136
Qy 282 ATGAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTCCAGTCTCTTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAsp 156
Qy 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCCGAGGG---GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGln 176
Qy 393 CGGGTTCAGCCCGCCAGGGCTCAGTGTGGCAGCGGGGGGCCCGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
Qy 453 GTGGACATCCCTTCGGCTCCTGCTGCGCCACCCAGTAGTGTGGTGCCCATTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
Qy 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTCGATAGG 572
```

```
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGGTGACCCCGCAGCGCTCAGCTGTGGCAGCGGGGCCCGCAGCAGAGAGCAAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGGCTCCTGGTGTGCCACCCAGCTATGCTGGTGGCTATTTGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGCTCCAGATAGAGCTATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACCGAGGTGAGCTGAAGAGCCATCAGTGTGCATCCACCCCTGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTGAAGATGATCTTGGAGATTATGATAAGAGCTAAGGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GTTACGAGGTTCCTCCGAGATCCTGGCCCATATAACTTTTAGGGCGTCTCATGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGACGAGCACTGAGAGGTAGAGCAAGTACGAGACACGACAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCCTTCCAGACCTTACCTCTTACACCTCGAGGAGCACCATCTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATGTTGACGGCCGAGCAGGAGAAATAATGAAGAAATTCGGGAGGCTATGAGAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTCCCATGAGC-----TCTCAGCTGATCCTGGCTGAGCACTGGCTGTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCACGAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTGCTCCTATAGCTCCTTTATGACGGTCCCGAGCAGGAGATGGTGCAAGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCAGCAGTGGCGCCATCATCGGCAAGAGGGGCGAGCAGCATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln 430
QY 1158 CTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCATGACCGCCAGAGCCCAATTCAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAGAGGAGAACTCTTTGGTCCCAAGCAGGACTGAGCTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CATCATGCTGTGCAGCATCAGCAGCTGGCGGGGTCTATTGGCAAAAGTGGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGACTTCAGAAATTTGACGCGAGCTGAGGTGGTAGTACCAAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACAGGTCTATCGTGAATAATCATCGGACATTTCTATGCCAGTCTAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValVallysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGCCCGCAGGTAAAGCAG--CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACACGCGCCAGCAGCAGCGAGGAAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
RESULT 4
US-09-850-716A-348
; Sequence 348, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-348
Alignment Scores:
Pred. No.: 1,52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 9 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-850-716A-348 (1-579)
QY 102 CGGTTAGGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTGTGTGGCTCAGTATGTACAGTAGAGAGTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGTGTGAATGTCACTATTCCAAACGGGAGCAGACCGGCAAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAACTGAAATGGCCACAGTTGGAGAACCATGCCCTGAAGTCTCTACATCCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGGTGACCCCGCAGCGCTCAGCTGTGGCAGCGGGGCCCGCAGCAGAGAGCAAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGGCTCCTGGTGTGCCACCCAGCTATGCTGGTGGCTATTTGGCAAG 512
```

QY 228 ACGGCTGTGTGAATGTCACCTATTCACCGGGAGCAGACCGAGCGAAGCCATCATGAAG 287
Db ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGGCCACAGTTGGAGAACCATCGCCTGAAGGTCTCTACATCCCGCATGAGCAG 347
Db LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
QY 348 ATAGCACAGGAGCCTCGAAGATGGCGCGGCGGCTTTGGCTCTCGGGGTTCAGCCCGC 407
Db IleThrGlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGCAGTGGCATCCCTT 467
Db GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 468 CGGCTCCTGTGGCCACCCAGTANGTGGTCCCATTTATGGCAAGGAGGGGCCACCATC 527
Db ArgLeuLeuValProThrGlnTyrValGlyAlaIleLeuGlyLysGluGlyAlaThrIle 218
QY 528 CGCAACATCACAAACAGACCCAGTCCCAAGATAGACGTGCATAGGAAGGAGAACCGAGGT 587
Db ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTGAAGAACCATCAGTGTGCATCCACCCCTGAGGCTGCTCTCCGCTTGAAG 647
Db AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 648 ATGATCTTGGAGATTATGCAAGAGAGGCTAAGACACCAAAACGGCTGACGAGTCC 707
Db MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTCGGCCCAATAAATCTTTGTAGGCGTCTCATTTGGCAAGGAAGCGGAAC 767
Db LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAGGTAGACAGATACCGAGCAAAATTCACCTCTCTCTGTTCCAGACCTT 827
Db LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTACACCCCTGAGAGACCATCATCTGGAAGGGGCCATCGAGAAATTTGTGACGG 887
Db ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
QY 888 GCCGAGCAGGAATAATAAGAAAGTTCCGGAGGCGCTATGAGAATGATGTGCTGCCATG 947
Db AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTACCTGATCCGCTGCTGAACTGGCTGTGTAGTCTTTCCAGCT 1001
Db SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1002 TCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGTCC 1061
Db SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTyrSerSer 398
QY 1062 TTTATGAGGCTCCCGAGCAGAGATGTTGAGGTGTTTATCCCGCCCGCAGGCGATGGGC 1121
Db PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGGCAAGAGGGGAGCAGACATCAAAACGCTCTCCCGTTTCCAGCGCCCTCC 1181
Db AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGACACCCGAAACCTGACTCCAAAGTTCGTATGTTATCATCACTCGA 1241
Db IleLysIleAlaProProGluThrProAspSerLysValArgMetValValIleThrGly 458
QY 1242 CCGCAGAGGCCCAATTAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGGAGAAC 1301
Db ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluLysAsn 478
QY 1302 TTCCTTGGTCCCAAGGAGGAAGTAGCTGGAGACCCACATACATCGTGTGCCAGCATCAGCA 1361

Db 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGCGCGGTTCATTGGCAAAGGTGGAACCGGTGAACGAGTTGCGAATTTGACGGCA 1421
Db AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGTGTGTAGTACCAAGAGACACCCCTCATGAGAACCGACCGAGTTCATCGTGA 1481
Db AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGACATTTCTATGTCAGTCAGATGGCTCAACCGAAGATCCGAGACATCTCTGGCC 1541
Db IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGTTTAAGCAGCAGCATCAAGAGGACAGAGTAAACCGGCCCGCAGGACCGAGGAAG 1598
Db GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577
RESULT 3
US-09-735-705-348
; Sequence 348, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-705-348
Alignment Scores:
Pred. No.: 1.52e-133 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 9 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-735-705-348 (1-579)
QY 102 CGGGTTAGGAGCGGAAATTCAAATCCGAAATATTCACCCCGAGCTCCGATGGGAAGTA 161
Db ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGACAGCCTGCTGCTCAGTATGGTACGTAGAGAACTGTGAGCACTGTGAGCAAGTGAACCCGAG 221
Db LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGCGCAGTGGTGAATGTCACTATTCCACCGGAGCAGACCCAGCAAGCCCATC 281
Db SerGluThrAlaValValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGGCCACAGTTCGAGAACCATCGCTGAAGGTCTCTACATCCCGCAT 341

```
Pred. No.: 1.09e-181 Length: 577
Score: 2509.00 Matches: 497
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 71.06% Indels: 2
DB: 15 Gaps: 1

US-09-270-437D-7 (1-1946) x US-10-313-986-500 (1-577)

QY 108 AGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGAGTACTGGAC 167
DB 79 ArgSerArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAsp 98
QY 168 AGCTCTGCTCAGTATGTACAGTAGAGAACCTGTGAGCAAGTGAACACCGAGGTGAG 227
DB 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
QY 228 ACGSCAGTGTGAATGTCACCTATTCCAAACGGGAGCAGACCGGCAAGCCATCATGAAG 287
DB 119 ThrAlaValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGCGCCACCATGTTGGAGAACCATGCGCTGAAAGGTCTCTACATCCCGATGAGCAG 347
DB 139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGln 158
QY 348 ATAGCAGGACCTGAGATGGCGCGCGAGGGGCTTTGGCTCTCGGGTCAAGCCCGC 407
DB 159 IleAlaGlnGlyProGluAsnGlyArgArgGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGGCTCACTGTGGCAGCGGGGGCCCCCAGCAGCAGCAAGTGGACATCCCGCTT 467
DB 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
QY 468 CGCTCTCTGTGCCACCCAGTATGTGGGTGCCATTATTGCGCAAGGAGGGGGCCACCATC 527
DB 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 528 CGCAGATCTCAAAACAGACCCAGTCCAGATAGACGTGATAGGAGGAGCGAGGT 587
DB 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTGAAAGGCATCAGTGTGCTCACTCCACCCCTCAGGCGTCTCCCTCGCTGTAAG 647
DB 239 AlaAlaGlnLysAlaIleSerValHisSerThrProGlnGlyCysSerSerAlaCysLys 258
QY 648 ATGATCTTGAGATTATGCATAAAGAGCTTAAGACACCAAAACGGCTGACGAGTTCCC 707
DB 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAGATCTGCGCCCATATAACTTTGTAGGGGCTCTCATTTGGCAAGGAGCGAAC 767
DB 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAAGGTAGAGCAAGATACCGAGACAAAATACCATCTCTCTCGTTGCAAGACCTT 827
DB 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCCATPCGAGAATTGTTCAGG 887
DB 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
QY 888 GCGCAGCAGGAATAATGACAGAAAGTTCGGAGGCGCTATGAGATGATGTGGTCCCATG 947
DB 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTACCTGTATCCCTGGCTGAACCTGGCTGTGTAGTCTTTTCCAGCT 1001
DB 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1002 TCATPCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGTCTCCCTATAGCTCC 1061
DB 379 SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTyrSerSer 398
```

```
1062 TTTATGACGGCTCCCGAGCAGGAGATGGTGCAAGTGTATATCCCGCCCGCAGGAGTGGC 1121
DB 399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCGCAGGAGGCGGAGCAGCATCAACACAGCTCTCCCGGTTTGCACGGCTCC 1181
DB 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCACTGGA 1241
DB 439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleIleThrGly 458
QY 1242 CCSCCAGAGGCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAAC 1301
DB 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeuLysGluGluAsn 478
QY 1302 TTCTTTGGTCCCAAGGAGGAAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
DB 479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGGCGGGTCAATTGCGCAAAGGTGGAATAACGGTGAACGAGTTGCAGAAATTTGACGGCA 1421
DB 499 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGGTGTAGTACCAAGAGACCAAGACCCCTGATGAGAACACAGGTCATCGTGAAA 1481
DB 519 AlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGGAAGATCCGAGACATCCTGGCC 1541
DB 539 IleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACAGGCCCCAGCGCAGGAGGAAG 1598
DB 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArgLys 577

RESULT 2
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-873-637-2

Alignment Scores:
Pred. No.: 2.54e-180 Length: 577
Score: 2491.00 Matches: 493
Percent Similarity: 99.00% Conservative: 1
Best Local Similarity: 98.80% Mismatches: 3
Query Match: 70.55% Indels: 2
DB: 9 Gaps: 1

US-09-270-437D-7 (1-1946) x US-09-873-637-2 (1-577)

QY 108 AGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGAGTACTGGAC 167
DB 79 ArgSerArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAsp 98
QY 168 AGCTCTGCTCAGTATGTACAGTAGAGAACCTGTGAGCAAGTGAACACCGAGGTGAG 227
DB 99 SerLeuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu 118
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 11:15:41 ; Search time 92.8527 Seconds
(without alignments)
13126.119 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtagcggaggctggg.....atttcctcaggtttaaaa 1946

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 2576884

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ n2p.model -DEV=xlsp
-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_23072004_095742_27560/app_query.fasta_1.5582
-DB=Published_applications AA -QPM=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITIS=bits -START=1 -END=-1 -MAIRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCOR=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODB=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09270437 @CGN 1 1 35 @runat 23072004 095742 27560
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/CTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

ALIGNMENTS

RESULT 1
US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-500

Alignment Scores:

1	2509	71.1	577	15	US-10-313-986-500	Sequence 500, App
2	2491	70.5	577	9	US-09-873-637-2	Sequence 2, Appli
3	1875	53.1	579	9	US-09-735-705-348	Sequence 348, App
4	1875	53.1	579	9	US-09-850-716A-348	Sequence 348, App
5	1875	53.1	579	9	US-09-897-778-348	Sequence 348, App
6	1875	53.1	579	9	US-09-897-778-446	Sequence 446, App
7	1875	53.1	579	9	US-09-897-778-449	Sequence 449, App
8	1875	53.1	579	12	US-10-007-700-348	Sequence 348, App
9	1875	53.1	579	12	US-10-007-700-446	Sequence 446, App
10	1875	53.1	579	12	US-10-007-700-449	Sequence 449, App
11	1875	53.1	579	14	US-10-117-982-348	Sequence 348, App
12	1875	53.1	579	14	US-10-117-982-446	Sequence 446, App
13	1875	53.1	579	14	US-10-117-982-449	Sequence 449, App
14	1875	53.1	579	14	US-10-117-982-480	Sequence 480, App
15	1875	53.1	579	15	US-10-313-986-348	Sequence 348, App
16	1875	53.1	579	15	US-10-313-986-446	Sequence 446, App
17	1875	53.1	579	15	US-10-313-986-449	Sequence 449, App
18	1875	53.1	579	15	US-10-313-986-480	Sequence 480, App
19	1875	53.1	586	9	US-09-850-716A-427	Sequence 427, App
20	1875	53.1	586	9	US-09-897-778-427	Sequence 427, App
21	1875	53.1	586	12	US-10-007-700-427	Sequence 427, App
22	1875	53.1	586	14	US-10-117-982-427	Sequence 427, App
23	1875	53.1	586	15	US-10-313-986-427	Sequence 427, App
24	1875	53.1	589	15	US-10-313-986-486	Sequence 486, App
25	1873	53.0	579	9	US-09-735-705-176	Sequence 176, App
26	1873	53.0	579	9	US-09-850-716A-176	Sequence 176, App
27	1873	53.0	579	9	US-09-897-778-176	Sequence 176, App
28	1873	53.0	579	10	US-09-466-396A-176	Sequence 176, App
29	1873	53.0	579	12	US-10-007-700-176	Sequence 176, App
30	1873	53.0	579	14	US-10-117-982-176	Sequence 176, App
31	1873	53.0	579	15	US-10-313-986-176	Sequence 176, App
32	1871	53.0	579	14	US-10-117-982-484	Sequence 484, App
33	1871	53.0	579	15	US-10-313-986-484	Sequence 484, App
34	1763.5	49.9	620	9	US-09-764-864-1116	Sequence 1116, Ap
35	1757.5	49.8	587	15	US-10-313-986-501	Sequence 501, App
36	1697	48.1	556	14	US-10-097-340-147	Sequence 147, App
37	1697	48.1	556	16	US-10-648-593-182	Sequence 182, App
38	1573.5	44.6	555	15	US-10-262-445-40	Sequence 40, Appl
39	1560	44.2	422	16	US-10-408-765A-2088	Sequence 2088, Ap
40	1258	34.6	261	9	US-09-764-864-1114	Sequence 1114, Ap
41	1228	34.8	250	9	US-09-764-864-1532	Sequence 1532, Ap
42	653	18.5	171	9	US-09-764-864-1119	Sequence 1119, Ap
43	510	14.4	192	9	US-09-764-864-1117	Sequence 1117, Ap
44	502	14.2	171	9	US-09-764-864-1536	Sequence 1536, Ap
45	361	10.2	81	14	US-10-117-982-476	Sequence 476, App

This Page Blank (uspto)

QY 699 ---GAGGTTCCCTGAAGATCTCGCCCATATAACTTTGTAGGCGTCTCATTTGGCAAG 755
Db 184 GlyAsnAlaValGlnGluLeuMetIleProAlaSerLysAlaGlyLeuValIleGlyLys 203
QY 756 GAAGGACGGAACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCATCTCTCCG 815
Db 204 GlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetValMet----- 221
QY 816 TTGCAAGACCTTACCTTTTACAACTCTGAGAGACC-----ATCACTGTG 860
Db 222 IleGlnAsp-----GlyProGlnAsnThrGlyAlaAspLysProLeuArgIle 237
QY 861 AAGGGGCGCATCGAAGATTGTTCAGGCGCGAGAGAAATAATGAAGAAATTCGGGAG 920
Db 238 ThrGlyAspProTyLysValGlnGlnAlaLysGluMetValLeuGluLeuIleArgAsp 257
QY 921 -----GCCTATGAGATGATGGCTGCCATGAGCTCTCACCTGATCCCT 965
Db 258 GlnGlyGlyPheArgGluValArgAsnGluTyfGly----- 269
QY 966 GGCCTGAACCTGGCTGTGTAGTCTTTTCCAGCTTCAATCCAGCGCAGTCCCGCGCT 1025
Db 269 ----- 269
QY 1026 CCCAGCAGCGTTACTGGGCTGCTCCTATAGTCTCTTTATCAGGCTCCCGAGCAGGAG 1085
Db 270 ---SerArgIleGlyGly-----AsnGlu 276
QY 1086 ATGGTGCAGGTGTTATCCCGCCCGCAGGAGTGGGCGCATCATCGCAAGAGGGGAG 1145
Db 277 GlyIleAspValProIleProArgPheAlaValGlyIleValIleGlyArgAsnGlyGlu 296
QY 1146 CACATCAACAGCTCTCCGGTTCGACGCGCTCCATCAAGATTGCACCCCGAA--- 1202
Db 297 MetIleLysLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspGly 316
QY 1203 ---ACACCTGACTCCAAAGTTCGTATGGTTATCATCATCGACCGCCAGAG---GCCCAA 1256
Db 317 ThrThrProGlu-----ArgIleAlaGlnIleThrGlyProProAspArgCysGln 333
QY 1257 TTCAGGCTCAG----- 1268
Db 334 HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353
QY 1269 -----GGAAGATCTATGCAAACTCAAGGAGGAGAACTTCTTGTGTCCTCAAG 1316
Db 354 ProGlyProGlyGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro 373
QY 1317 GAGGAAGTGAAGTGGAGCCACATACGTGTGCCAGATCATCAGCTGGCCGGGTCAAT 1376
Db 374 GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle 392
QY 1377 GCGAAGTGGAAACCGTGAACGTTGACAGTTGACAGATTTACCGCAGCTGAGGTGTAGTA 1436
Db 393 GlyLysGlyGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeu 412
QY 1437 CCAAGAGACACGACCTGATGAGAACACCGAGTC---ATCGTGAATAATCATCGGACAT 1493
Db 413 GlnArgAsnProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432
QY 1494 -----TTCTATCCAGTCAAGTGGCTCAACGGAAGAT-----CCGAGACAT 1534
Db 433 ProGlnGlnIleAspTyfAlaArgGlnLeuIleGluLysIleGlyGlyProValAsn 452
QY 1535 CTT---GCCGAGTTTAACGACGATCAGAAAGGACAGAGTAACAGGCCGAGGACG 1591
Db 453 ProLeuGlyProProValProHisGlyProHisGlyVal-----ProGlyPro--HisG 470
QY 1592 GA-----GGAAGTGACCGCCCTC-----CCTGTCCCTTNGA 1624
Db 470 IyProProGlyProProGlyProGlyThrProMetGlyProTyfAsnProAlaProTyfA 490

QY 1625 GTCCAGGA----- 1632
Db 490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyfAlaProGlnGlyT 510
QY 1633 -----CAACAACGGGCGAGAAATCGAGAGTGTGCTCTCCCGG 1669
Db 510 rpGlyAsnAlaTyfProHisTyfGlnGlnAlaProProAsp-----ProA 526
QY 1670 CAGGCTCGAGATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGC 1729
Db 526 laLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyf-----TyrAla 542
QY 1730 CACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTNTCAGCCCAACACCCAC 1789
Db 543 HisTyfTyfGlnGlnAla-----GlnProProAlaAla 555
QY 1790 CCAATTGGC---CCAACACTGTNTGCCCTCCGGGTGTGAGAAATTNTAGCGCAGGCAC 1846
Db 556 ProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGlyAsp----- 570
QY 1847 TTTTAAACGTGATTGTTTAAAGAAGCTCTCAGGCCCCCAAGAGGGTGTGATCAGCACC 1906
Db 571 ---GlnGlnAsnProAlaProAlaGlyGlnValAspTyfThr 593
QY 1907 -----TCAGTGGGAAGAAATAATAATTTCTTTCAGT 1939
Db 584 LysAlaTrpGluGluTyfTyfLysLysMetGlyGlnAlaValProAlaProThrGly 602

Search completed: July 23, 2004, 11:19:34
Job time : 44.4913 secs


```
Db 413 GlnArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThr 432
QY 1494 -----TTCATGCGCAGTCAGATCGCTCAACGGGAAGAT-----CCGAGACAT 1534
Db 433 ProGlnGlnIleAspTyrAlaArgGlnLeuIleGlnGlyLysIleGlyGlyProValAsn 452
QY 1535 CCT---GGCCCGAGTTAAGCAGCAGCATCAGAAGGGACAGAGTAACAGGCCCGCAGGCAAG 1591
Db 453 ProLeuGlyProValProHisGlyProHisGlyVal-----ProGlyPro-HisG 470
QY 1592 GA-----GGAGTCACCGCCCTC-----CCTGTCCCTTNGA 1624
Db 470 lYProProGlyProProGlyProGlyThrProMetGlyProTyrAsnProAlaProTyrA 490
QY 1625 GTCCAGGA----- 1632
Db 490 snProGlyProProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyT 510
QY 1633 -----CAACACGGGAGAGAAATCGAGAGTGCTCTCCCGG 1669
Db 510 rpGlyAsnAlaTyrProHisTrpGlnGlnGlnAlaProProAsp-----ProA 526
QY 1670 CAGSCCTGAGATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGC 1729
Db 526 lAlysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAlaTyr-----TyrAla 542
QY 1730 CACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTNTCAGGCCCAACACCCAC 1789
Db 543 HisTyrTyrGlnGlnGlnAla-----GlnProProProAlaAla 555
QY 1790 CCAATGGC---CCACACTGTNTGCCCTCCGGGTGTCAAGAAATNTAGCGGAAGCAC 1846
Db 556 ProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGlyAsp----- 570
QY 1847 TTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGGCCCGCCACCAAGAGGGTGGATCACAC 1906
Db 571 -----GlnGlnAsnProAlaProAlaGlyGlnValAspTyrThr 583
QY 1907 -----TCAGTGGGAAGAAAAATAAAATTTCTCTCAGGT 1939
Db 584 LysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaValProAlaProThrGly 602

RESULT 15
PCT-US94-01782-2
; Sequence 2, Application PC/TUS9401782
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
; APPLICANT: HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: NOVEL FUSE BINDING
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01782
; FILING DATE: 22-FEB-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/021,608
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHEetical: No
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE: HL60
; ORGANELLE:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile
; PCT-US94-01782-2

Alignment Scores:
Pred. No.: 1,83e-13 Length: 644
Score: 245.00 Matches: 148
Percent Similarity: 36.46% Conservative: 93
Best Local Similarity: 22.39% Mismatches: 219
Query Match: 6.94% Indels: 202
DB: 5 Gaps: 33

US-09-270-437d-7 (1-1946) x PCT-US94-01782-2 (1-644)
QY 309 AACATGCTCCCTGAAGGTCTCTACATCCCGAGTAGACAGATACACAG-----GGACCT 362
Db 28 AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGlyGly 47
QY 363 GAGATGGG-----CGCCGAGGGGGTCTCGGGTCTCGGGTCTAGCCCGCCAGGGC 413
Db 48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrTyrGlyGlyGlnLysArg----- 65
QY 414 TCACCTGTGGCAGCGGG----- 431
Db 66 ---ProLeuGluAspGlyAspGlnProAspAlaLysLysValAlaProGlnAsnAspSer 84
QY 432 -----GCCCGACCAAGCAGCAGCAAGTGGACATCCCTCT-----CGG 470
Db 85 PheGlyThrGlnLeuProProMetHisGlnGlnGlnSerArgSerValMetThrGluGlu 104
QY 471 CTCCTGTGTGCCACCCAGTATGCGGTGCCATTATTGGCAAGGAGGGGGCCACATCCGC 530
Db 105 TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGlnIleSer 124
QY 531 AACATCACAACAGACCCAGTCCAAGATAGACGTAGTAGGAAGAGAACCGAGGTGCA 590
Db 125 ArgIleGlnGlnSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeu 143
QY 591 GCTGAAAAGCCATCAGTGTGCACTCCACCTGAGGGGTCTCTCCGCTTTAAGATG 650
Db 144 ProGluArgSerCys***LeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu 163
QY 651 ATCTTGGAGATTATGCATAAA-----GAGGCTAAGGACACCAACACGGGTGAC----- 698
Db 164 LeuAspGlnIleValGluLysGlyArgProAlaProGlyPheHisIleGlyAspGlyPro 183
```

APPLICANT: ROBERT C., AND AVIGAN, MARK I.
 TITLE OF INVENTION: NOVEL FUSE BINDING
 TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
 NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/726,160

FILING DATE: 04-OCT-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/021,608

FILING DATE: 22-FEB-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: WILLIAM S. FEILER

REGISTRATION NUMBER: 26,728

REFERENCE/DOCKET NUMBER: 2026-4063US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 644

TYPE: Amino Acid

STRANDEDNESS: Single

TOPOLOGY: Unknown

MOLECULE TYPE: Peptide/Protein

HYPOTHETICAL: No

ORGANISM: Human

CELL LINE: HL60

FEATURE:

OTHER INFORMATION:

OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile

US-08-726-160-2

Alignment Scores:

Pred. No.: 1.83e-13 Length: 644

Score: 245.00 Matches: 148

Percent Similarity: 36.46% Conservative: 93

Best Local Similarity: 22.39% Mismatches: 219

Query Match: 6.94% Indels: 202

DB: 1 Gaps: 33

US-09-270-437D-7 (1-1946) x US-08-726-160-2 (1-644)

QY 309 AACCATGCCCTGAAGTCTCTACATCCCGATGAGCAGATAGCACAG-----GGACCT 362

Db 28 AsnAspAlaPheLysAspAlaLeuGlnArgAlaArgGlnIleAlaAlaLysIleGlyGly 47

QY 363 GAGATGGG-----CGCCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCGAGGC 413

Db 48 AspAlaGlyThrSerLeuAsnSerAsnAspTyrGlyTyrGlyGlnLysArg----- 65

QY 414 TCACCTGTGGCGCGGG----- 431

Db 66 ---ProLeuGlnAspGlyAspGlnProAspAlaLysLysValalaProGlnAsnAspSer 84

QY 432 -----GCCCCAGCAGCAGCAGTGGACATCCCCCT-----CGG 470

Db 85 PheGlyThrGlnLeuProProMetHisGlnGlnSerArgSerValMetThrGluGlu 104

471 CTCCTGGTCCCAACCCAGTATGTGGTGCCTATTATGGCAAGGAGGGGCCACCATCCGC 530
 105 TyrLysValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGlnIleSer 124
 531 AACATCAAAACAGACAGCCAGTCCAAGATAGACGTGCATAGGAAGAGAAACAGAGTGA 590
 125 ArgIleGlnGlnGlnSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeu 143
 591 GCTGAAAAAGCATCAGTGTGCATCCACCCCTGAGGGCTGCTCTCCGCTTGTAGATG 650
 144 ProGluArgSerCys**LeuThrGlyThrProGluSerValGlnSerAlaLysArgLeu 163
 651 ATCTTGGAGATATGCATAA-----GAGCTAAGGACACCAACACCGCTGAC----- 698
 164 LeuAspGlnIleValGluLysGlyArgProAlaProGlyPheHisHisGlyAspGlyPro 183
 699 --GAGGTTCCCTGAAGATCTGCCCCATAATACTTTGTAGGCGCTCTCATTTGGCAAG 755
 184 GlyAsnAlaValGlnGluIleMetIleProAlaSerLysAlaGlyLeuValIleGlyLys 203
 756 GAAGACGGAACCTGAAGAGGTAGACCAAGATCCGAGACAAATAATCACCATCTCTCG 815
 204 GlyGlyGluThrIleLysGlnLeuGlnArgAlaGlyValLysMetValMet----- 221
 816 TTGCAAGACCTTACCTTTACACCTGAGAGGACC-----ATCACTGTG 860
 222 IleGlnAsp-----GlyProGlnAsnThrGlyAlaAspLysProLeuArgIle 237
 861 AAGGGGGCCATCGAGATTTGTTCAGGGCCGAGCAGAGAAATAATGAAGAAATTCGGGAG 920
 238 ThrGlyAspProTyrLysValGlnGlnAlaLysGluMetValLeuGluLeuIleArgAsp 257
 921 -----GCCTATGAGATGATGTGGCTGCCATGAGCTCTCACCTGATCCCT 965
 258 GlnGlyGlyPheArgGluValArgAsnGluTyrGly----- 269
 966 GGCCTGAACCTGGCTGTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCCT 1025
 269 ----- 269
 1026 CCCAGCAGGTACTCGGGCTGCTCCTATAGCTCTTTATGCAGGCTCCCGAGCAGGAG 1085
 270 ---SerArgIleGlyGly-----AsnGlu 276
 1086 ATGTGTCAGGTGTTTATCCCGCCAGCAGTGGCGGCATCATCGCAAGAGGGGCGAG 1145
 277 GlyIleAspValProIleProArgPheAlaValGlyIleValIleGlyArgAsnGlyGlu 296
 1146 CACATCAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGATTGCAACCCCGAA--- 1202
 297 MetIleLysLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspAspGly 316
 1203 ---ACACCTGACTCCAAAGTTTCGTATGTTATCATCATCTGGACCGCCAGAG---GCCAA 1256
 317 ThrThrProGlu-----ArgIleAlaGlnIleThrGlyProProAspArgCysGln 333
 1257 TTCAAGGCTCAG----- 1268
 334 HisAlaAlaGluIleIleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGly 353
 1269 -----GGAGAGATCTATGGCAACTCAAGAGGAGAGACTCTTTTGGTCCCAAG 1316
 354 ProGlyProGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProPro 373
 1317 GAGGAAGTGAAGCTCGAGACCCACATAGCTGTGCCAGCATCAGACGCTGGCGGGTCA 1376
 374 GlyGlyLeuGln---GluPheAsnPheIleValProThrGlyLysThrGlyLeuIleIle 392
 1377 GCAAGGTGGAAAAACGGTGAACCGATTGCGAGATTTCACGCGCAGCTGAGGTGAGTA 1436
 393 GlyLysGlyGlyGluThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGluLeu 412
 1437 CCAAGAGACCGACCCCTCATGAGAACGACCGAGTGC---ATCGTGAATAATCATCGACAT 1493

Db 386 LysThrGlyLeullelleGlyLysGlyGlyGluThrIleYsserIleSerGlnGlnSer 405
QY 1419 GCAGCTGAGGTGTAGTACCAAGACAGACCCCTGATGAGACACGAGTC---ATC 1475
Db 406 GlyAlaArgIleGluGlnArgAsnProProGlnGlnAlaAspProAsnMetLysLeu 425
QY 1476 GTGAAATCATCGACAT-----TTCTATGCCAGTCCAGTCCCTCAAGGAAG 1523
Db 426 PheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeulleGluGluLys 445
QY 1524 AT-----CGGACACATCCT---GGCCAGGTTTAAGCAGCAGCATCAGAGGACACAG 1573
Db 446 IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal--- 464
QY 1574 TAACAGCCCGCCAGCAGGA-----GGAAGTGACAGCAGCCCTC----- 1611
Db 465 ---ProGlyPro--HisGlyProProGlyProProGlyProGlyThrProMetGlyProT 483
QY 1612 ---CCTGTCCTTNGAGTCCAGA----- 1632
Db 483 YrAsnProAlaProTyrAsnProGlyProProGlyProAlaProHisGlyProProAlaP 503
QY 1633 -----CAACAAGCGGCAGAAATCG 1651
Db 503 roTyrAlaProGlnGlyTyrGlyAsnAlaTyrProHisTyrGlnGlnGlnAlaProProA 523
QY 1652 AGAGTGTGCTCTCCCGCGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGCGCGG 1711
Db 523 sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTyrAlaAla 538
QY 1712 CTGTAGATCAGTTTGCCCATCTGATGAGAAAGATGTCAGTGAGGAACCTGATCTN 1771
Db 539 Tyr-----TyrAlaHisTyrTyrGlnGlnGlnAla----- 548
QY 1772 TCAGCCCCAACACCCCAATGGC---CCAACACTGTNTGCCCTCGGCGTGTCTAGA 1828
Db 549 GlnProProAlaAlaProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGly 568
QY 1829 AATTNTAGCGCAGGCACTTTTAAACGTGGATTGTTTAAAGAAGCTCTCCAGGCCCCACC 1888
Db 569 Asp-----GlnGlnAsnProAlaProAla 576
QY 1889 AAGAGGTGATCATCAC-----TCAGTGGGAAGAAAATA 1924
Db 577 GlyGlnValAspTyrThrLysAlaTrpGluGluTyrTyrLysLysMetGlyGlnAlaVal 596
QY 1925 AAATTTCCTTCAGT 1939
Db 597 ProAlaProThrGly 601

RESULT 12
US-09-261-855-18
; Sequence 18, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-18
Alignment Scores:
Pred. No.: 4,67e-14 Length: 49
Score: 245.00 Matches: 49
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.94% Indels: 0
DB: 3 Gaps: 0
US-09-270-437D-7 (1-1946) x US-09-261-855-18 (1-49)
QY 714 ATCCCTGCCCAATAAATCTTTGTAGGGCTCTCATTGGCAAGGACGAACTGAAG 773
Db 1 IleLeuAlaHisAsnAsnPheValGlyArgLeulleGlyLysGluGlyArgAsnLeuLys 20
QY 774 AAGCTAGACGAAATACCGAGACAAAAATACCATCTCTCTCGTTGCAAGACCTTACCCTT 833
Db 21 LysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeuThrLeu 40
QY 834 TACAACCTGTAGAGACCATCAGTGTG 860
Db 41 TyrAsnProGluArgThrIleThrVal 49

RESULT 13
US-08-021-608D-2
; Sequence 2, Application US/08021608D
; Patent No. 5580760
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile
US-08-021-608D-2
Alignment Scores:
Pred. No.: 1.83e-13 Length: 644

OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
 US-08-726-160-10

Alignment Scores:

Pred. No.: 8,81e-14 Length: 643
 Score: 248.50 Matches: 140
 Percent Similarity: 37.80% Conservative: 97
 Best Local Similarity: 22.33% Mismatches: 202
 Query Match: 7.04% Indels: 189
 DB: 1 Gaps: 32

US-09-270-437D-7 (1-1946) x US-08-726-160-10 (1-643)

```

Qy 336 CCCGAT---GAGCAGATAGCAGGACCTGAGATGGCGCCGAGGGGGCTTGGCTCT 392
Db 73 ProAspAlaLysIysValAla-----ProGlnAsn-----AspSerPheGlyThr 87
Qy 393 CGGGGTGAGCCGCCAGGGCTCACCTGTGCGAGGGGGGGCCAGCCGACGACGACAA 452
Db 88 GlnLeuProProMetHis-----GlnGlnGlnArg 97
Qy 453 GTGGACATCCCCCTCGCTCCTGTGTCGCCACCCAGTATGTTGGTGCCATTATTGGCAAG 512
Db 98 SerValMetThrGluGluTyrLysValProAspGlyMetValGlyPheIleIleGlyArg 117
Qy 513 GAGGGGGCCACATCCGCAACATCACAAACAGACCCAGTCCACAGATAGACGTGCATAGG 572
Db 118 GlyGluGlnIleSerArgIleGlnGlnGlnSerGlyCysLysIleGlnIle---Ala 136
Qy 573 AAGGAGAACGGGTGAGCTGAGTGAAGAACCATCATGTGTCCTCCACCCCTGAGGGTGC 632
Db 137 ProAspSerGlyGlyLeuProGluArgSerCys***LeuThrGlyThrProGluSerVal 156
Qy 633 TCCTCCGCTTGAAGATGATCTTGAGATTATGCAATAA-----GAGGCTAAGGACACC 686
Db 157 GlnSerAlaLysArgLeuLeuAspGlnIleValGluLysGlyArgProAlaProGlyPhe 176
Qy 687 AAAACGGGCTGAC-----GAGGTTCCCTCGAAGATCCTGGCCCAATAAATCTTGTGA 737
Db 177 HisHisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla 196
Qy 738 GGGGGTCTCATTTGGCAGGAAGGAGGAACTGTGAAGAGGTAGACCAAGATACCGAGACA 797
Db 197 GlyLeuValIleGlyLysGlyGlyLeuThrIleLysGlnLeuGlnGluArgAlaGlyVal 216
Qy 798 AAATATCACCATCTCCTCTGTCAGACACCTTACCCTTTTACACCCCTGAGAGGACC----- 851
Db 217 LysMetValMet-----IleGlnAsp-----GlyProGlnAsnThrGlyAla 230
Qy 852 -----ATCAGTGTGAAGGGGCCCATCGAGAAATTTGTCAGGGCCGAGCAGGAAATA 902
Db 231 AspLysProLeuArgIleThrGlyAspProTyrLysValGlnGlnAlaLysGluMetVal 250
Qy 903 ATGAAGAAGTTCGGAG-----GCCTATGAGATGATGTGCTGCCATG 947
Db 251 LeuGlnLeuIleArgAspGlnGlyGlyPheArgGluValArgAsnGluTyrGly----- 268
Qy 948 AGCTCTCAGCTGATCCCTGGCTGAACCTGGCTGCTGTAGGTCTTTCCAGCTTCATCC 1007
Db 268 ----- 268
Qy 1008 AGCGAGTCCGCGCGCTCCAGCAGCGTTCAGTGGGTGCTCCTATAGTCTCTTTATG 1067
Db 269 -----SerArgIleGlyGly----- 273
Qy 1068 CAGGCTCCGAGCAGGAGATGTGTCAGGTGTTATCCCGCGCCAGCAGTGGCGCCATC 1127
Db 274 -----AsnGluGlyIleAspValProIleProArgPheAlaValGlyIleVal 289
Qy 1128 ATCGGCAAGAGGGGCACATCAACAGCTCTCCCGTGTGGCCAGCGCTCCATCAAG 1187
Db 290 IleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspAlaGlyValArgIleGln 309

```

RESULT 11

PCT-US94-01782-10
 ; Sequence 10, Application PC/TUS9401782
 ; GENERAL INFORMATION:

```

Qy 1188 ATTGCACCCCGGAA-----ACACCTGACTCCAAGTTTCGTATGTTTATCATCATCGTA 1241
Db 310 PheLysProAspAspGlyThrThrProGlu-----ArgIleAlaGlnIleThrGly 326
Qy 1242 CCGCCAGAG--GCCCAATTCAAGGCTCAG----- 1268
Db 327 ProProAspArgCysGlnHisAlaAlaGluIleIleThrAspLeuLeuArgSerValGln 346
Qy 1269 -----GGAGAATCTATGCGAATCTGCAAACTCAAGGAGAG 1298
Db 347 AlaGlyAsnProGlyGlyProGlyProGlyArgGlyArgGlyArgGlyGlyGlyAsn 366
Qy 1299 AACTTCTTTGTCCTCCAGGAGGAAGTGAAGTGGAGCCCATACATACGTGTCGACATCA 1358
Db 367 TrpAsnMetGlyProProGlyGlyLeuGln---GluPheAsnPheIleValProThrGly 385
Qy 1359 GCAGCTGCGCGGTTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTCAGAAATTTGACG 1418
Db 386 LysThrGlyLeuIleIleGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 405
Qy 1419 GCAGCTGAGGTGGTGTAGTACCAAGAGACAGACCCCTGTATGAGACGACGAGTGC---ATC 1475
Db 406 GlyAlaArgIleGluLeuGlnArgAsnProProAsnAlaAspProAsnMetLysLeu 425
Qy 1476 GTGAAATCATCGACAT-----TTCTATGCCAGTCAGATCGGTCAACGGAG 1523
Db 426 PheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeuIleGluLys 445
Qy 1524 AT-----CGAGACATCCT---GGCCACAGTTAAGCAGCAGCATCAGAAAGGACAGAG 1573
Db 446 IleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHisGlyVal--- 464
Qy 1574 TAACGAGCCCGACGACGCA-----GGAAGTGACGACGCCCCCTC----- 1611
Db 465 ---ProGlyPro--HisGlyProProGlyProGlyProGlyThrProMetGlyProT 483
Qy 1612 ---CCTGTCCCTTNGAGTCCAGGA----- 1632
Db 483 TyrAsnProAlaProTyrAsnProGlyProGlyProGlyProAlaProHisGlyProProAlaP 503
Qy 1633 -----CAACAACGGCGCAGAAATCG 1651
Db 503 roTyrAlaProGlnGlyTrpGlyAsnAlaTyrProHisTrpGlnGlnAlaProProAla 523
Qy 1652 AGAGTGTCTCTCCCGCGAGCCCTCAGAAATGAGTGGAAATCCGGGACACNTGGCCGGG 1711
Db 523 sp-----ProAlaLysAlaGlyThrAspProAsnSerAla-AlaTrpAlaAla 538
Qy 1712 CTGTAGATCAGTTTGGCCACTTGATTGAGAAAGATGTTCCAGTGAAGACCTGATCTN 1771
Db 539 Tyr-----TyrAlaHisTyrTyrGlnGlnAla----- 548
Qy 1772 TCAGCCCAACACACCCCAATTTGGC---CCACACTGTNTGCCCTCGGGGTGTTCAGA 1828
Db 549 GlnProProProAlaAlaProAlaGlyAlaProThrThrThrGlnThrAsnGlyGlnGly 568
Qy 1829 AATTNTAGCGCAAGCACTTTTAAACGTGGATGTTTAAAGAGCTCTCCAGGCCCCACC 1888
Db 569 Asp-----GlnGlnAsnProAlaProAla 576
Qy 1889 AAGAGGTGGATCACC-----TCAGTGGGAGAAAAATA 1924
Db 577 GlyGlnValAspTyrThrLysAlaTrpGluTyrTyrLysLysMetGlyGlnAlaVal 596
Qy 1925 AAATTTCTCTCAGGT 1939
Db 597 ProAlaProThrGly 601

```



```

QY 873 GAGAAATGTTGCGAGGCGGAGGAGGAAATATGAGAAAGTTCCGGAGGCGCTATGAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluGluMetLysLysLysLysLysLysLysLysLysLys 353
QY 933 GATGCTGCTGCATGAGC-----TCTCACTGATCCCTGCGCTGAACTGCTGCTGTA 986
Db 354 AspLeuAlaSerMetAsnLeuGlnAlaHisLeuLeuProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCGAGCTTCATCCAGCGCAGTCCCGCGC-----CCTCCAGCAGCGGT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGCTGCTCCCTATAGCTCTTATGAGCGCTCCGAGCAGGAGATGCTGAGGTG 1097
Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY 1098 TTTATCCCGCCAGGAGTGGCGCATATCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1157
Db 411 PheLeuProAlaLeuSerValGlyAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 430
QY 1158 CTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACCCAGGAGGAGGAGGAGGAGGAG 1217
Db 431 LeuSerArgPheAlaGlyAlaSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 450
QY 1218 GTTCGTATGTTATCATCATCTGAGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1277
Db 451 ValArgMetValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGCAAACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCCAGCATACGAGCTGCGCGGTTCATTGGCAAAAGGTGGAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaGlyAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAAGATTGACGCGAGCTGAGTGTGTAGTACCAAGAGGAGGAGGAGGAG 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGACGACCGAGGTCATCGTGAATATCATCGACATTTCTATGCCAGTCAGATGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheThrAlaCysGlnValAlaGln 550
QY 1518 CGAAAGATCCGAGCATCTCGCCCGCAGGTAAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACGAGCGCCGAGCAGGAGGAAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579

```

RESULT 9

```

US-08-021-608D-10
; Sequence 10, Application US/08021608D
; Patent No. 5580760
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
; US-08-021-608D-10

```

```

Alignment Scores:
Pred. No.: 8,81e-14 Length: 643
Score: 248.50 Matches: 140
Percent Similarity: 37.80% Conservative: 97
Best Local Similarity: 22.33% Mismatches: 202
Query Match: 7.04% Indels: 189
DB: 1 Gaps: 32

```

```

US-09-270-437D-7 (1-1946) x US-08-021-608D-10 (1-643)
QY 336 CCCGAT---GAGCAGATAGCAGGAGCTGAGAAATGGCGCGCGAGGGGCTTGGCTCT 392
Db 73 ProAspAlaLysLysValAla-----ProGlnAsn-----AspSerPheGlyThr 87
QY 393 CGGGGTCAAGCCCGCCAGGCTCACCTGTGGCAGCGGGGCCGCCAGCCAGCAGCAAG 452
Db 88 GlnLeuProProMetHis-----GlnGlnGlnArg 97
QY 453 GTGGACATCCCTTCGGCTCTGTGTGCCACCCAGTATGTGGTGCCATTTATGGCAAG 512
Db 98 SerValMetThrGluGluTyrLysValProAspGlyMetValGlyPheIleIleGlyArg 117
QY 513 GAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCCAAGATAGACCTGCATAGG 572
Db 118 GlyGlyGluGlnIleSerArgIleGlnGlnGluSerGlyCysLysIleGlnIle---Ala 136
QY 573 AAGGAGAACCGCAGTGCAGCTGCAAAAAGCCATCAGTGTGTCACCTCCACCCCTGAGGCTGC 632
Db 137 ProAspSerGlyGlyLeuProGluArgSerCys**LeuThrGlyThrProGluSerVal 156
QY 633 TCCTCCGCTTGAAGATGATCTTGGAGATATTGCATAAA-----GAGGCTAAGGACACC 686
Db 157 GlnSerAlaLysArgLeuAspGlnIleValGluLysGlyArgProAlaProGlyPhe 176
QY 687 AAAACGGCTGAC-----GAGGCTCCCTGGAAGATCTCGGCCCATATAACTTTGTA 737
Db 177 HisHisGlyAspGlyProGlyAsnAlaValGlnGluIleMetIleProAlaSerLysAla 196
QY 738 GGGCGTCTCTATTGGCAAGGAGGAGCGAACTTGAAGAAGGTAGAGCAAGATACCGGAGCA 797
Db 197 GlyLeuValIleGlyLysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyVal 216

```

```
QY 813 TCCTGTCAGACCTTACCCCTTTACACCTCGAGAGACCATCATCTGTGAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTTGTCAGGGCGGACAGCAATATAGAAAGTTTCGGAGGCTATGAGAAT 932
Db 334 GluThrCysAlaLysAlaGluGluGluMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGCTGCATGAGC-----TCTCACTGATCCCTGGCCCTGAACCTGGCTGTCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCATGCCGCG-----CTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCGAGCAGGAGATGTCAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY 1098 TTATATCCCGCCAGGAGTGGGGCCCATCATCGGCAAGAGAGGGCGCGCATCAAAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACCCCGCAACACCTGCTCCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTGCTATGTTATCATCTGACCGCCAGCCGAGCCCAATTCAAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAACTCAAGGAGGAGAACTTCTTGGTCCCAAGAGAGAACTGAAGCTGGAGAC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CATACAGCTGTCAGCAGCTCAGCAGCTGCGCGGTCAATTCGCAAGGTTGGAAGACGGTG 1397
Db 491 HisIleArgValProSerPheAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGAGAAATTTGACGGCAGCTGAGTGTAGTACCAAGAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGACGACGAGTCTCGTGAATCATCGGACATTCATGCGAGTCAAGTGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGAGACATCTCGCCAGCTGAGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACGAGCCGAGGACGAGGAG 1598
Db 571 SerGlyProGlnSerArgLys 579
```

RESULT 8

```
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangor, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
```

```
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-606-421B-176

Alignment Scores:
Pred. No.: 8,4e-161 Length: 579
Score: 1873.00 Matches: 380
Percent Similarity: 84.48% Conservative: 50
Best Local Similarity: 74.66% Mismatches: 63
Query Match: 53.04% Indels: 16
DB: 4 Gaps: 9
```

US-09-270-437D-7 (1-1946) x US-09-606-421B-176 (1-579)

```
QY 102 CGGTTAGAGCGCGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGAAAGTA 161
Db 77 ArgGluArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTyrGluVal 96
QY 162 CTGGACAGCCCTGCTGGCTCAGTATGTGACAGTAGAGAACTGTGAGCAAGTGAACACGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGAGCGCAGTGTGATGTCACCTATTCACACGGGAGCAGACCGAGCCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCCACCCAGCTTGGAGAACCATCGCTCAAGGCTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgLysArgGlyLeuGlyGln 176
QY 393 CGGGGTGAGCGCGCGCGCGCTCACCTGTGGCAGCGGGGCGCCAGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGACATCCCTCTCGCTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCCAAGATAGACGTGCTAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGGTGCGCTGAAAGCCATCAGTGTGCACCTCCACCTCCAGGGGTGC 632
Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCGCTGTGAAGATGATCTTGGAGATTATGCATATAAGAGGCTAAGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACAGAGTTCCTCGCTGAGATCCCTGGGCGCATATAACTTTGTAGGGGTCTCANTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGAGCGAACTGTAAGAGGTAGAGAGATACCGAGACACAAATACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTSCAAGACCTTACCCCTTTTACACCTCGAGAGGACCATCACTGTGAAGGGGCGCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
```



```
Db      214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY      573  AAGGAGACGCGAGTGGAGCTGAAAGACCATCAGTGTGCATCTCACCCCTGAGGGCTGC 632
Db      234  LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY      633  TCCCTCCGCTTGAAGATGATCTCGAGATTATGCTAAAGAGGCTTAAGGACACCAACG 692
Db      254  SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspLysPhe 273
QY      693  GCTGACGAGTCCCTCAAGATCCTGCGCCATATTAATCTTTAGGGCGTCTCATTCGC 752
Db      274  ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY      753  AAGGAAGGACGACCTGGAAGAGGTAGACCAATACCGAGACAAAATCACCATCTCC 812
Db      294  LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY      813  TCCTGTCAAGACCTTACCTTTACAACCTCAGAGGACCATCACTGTGAAGGGGCCATC 872
Db      314  ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY      932  GAGATTGTTCAGGCGCAGCAGCAATAATGAAGAAAGTTCGGAGGCGCTATGAGAA 932
Db      334  GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY      933  GATGTGGCTCCATGAGC-----TCTCACCCTGATCCCTGGCCTGAACCTGGCTGTGA 986
Db      354  AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY      987  GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CTCCACGAGCAGGTT 1037
Db      374  GlyLeuPheProProThrSerGlyMetProProProProThrSerGlyProProSerAlaMet 393
QY      1038  ACTGGGGCTGCTCCTTAGCTCCTTTATGAGCTCCCGAGCAGGAGATGTTGCAGGTG 1097
Db      394  Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisGln 410
QY      1098  TTTATCCCGCCGAGGCGGCGCCATCATCGGCAAGAGGGGCGACATCAACAG 1157
Db      411  PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY      1158  CTCTCCGGTTGTCAGCGCTCCATCAAGATTGCACCCCGAAACCTGTACTCCAAA 1217
Db      431  LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY      1218  GTTCGTATGTTATCATCTGACCGCCAGAGCCCAATCAAGGCTCAGGGAAGATC 1277
Db      451  ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY      1278  TATGGCAACTCAAGGAGGAGACTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACC 1337
Db      471  TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY      1338  CACATACGTGTCAGCATCAGCAGCTGCGCGGTCTATTGGCAAAAGTGAAGGAAACGGTG 1397
Db      491  HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY      1398  AACGAGTTCAGAAATTGACGGCAGCTGAGGTGGTAGTACCAAGACACGACCCCTGAT 1457
Db      511  AsnGluLeuGlnAsnLeuSerAlaGluValValProArgAspGlnThrProAsp 530
QY      1458  GAGAACGACCATCTATCGTGAATCATCGGACATTTCTATCCAGTCAAGTGGCTCA 1517
Db      531  GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY      1518  CGAAGATCCGACATCTCTGGCCCGAGTTTAAGCAG---CAGCATCAGAAAGGA---CAG 1571
Db      551  ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY      1572  AGTAACGAGCCCGCAGGACGAGGAG 1598
Db      571  SerGlyProProGlnSerArgLys 579
```

```
RESULT 6
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.:      8.4e-161      Length:      579
Score:          1873.00      Matches:     380
Percent Similarity: 84.48%      Conservative: 50
Best Local Similarity: 74.66%      Mismatches:  63
Query Match:      53.04%      Indels:      16
DB:               4          Gaps:         9
```

US-09-270-437d-7 (1-1946) x US-09-480-884A-176 (1-579)

```
QY      102  CGGTTAGAGCCGCGAAATTCAAATCCGAAATATTCACCCAGCTCCGATCGGAAGTA 161
Db      77  ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnThrProGluVal 96
QY      162  CTGACAGCGCTGCTGGCTCAGTATGTGTACAGTACAGAACTGTGAGCAAGTGAACACCGAG 221
Db      97  LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY      222  AGTGAGACGCGCAGTGTGTATGTACCTATTCACCGGAGGAGACAGGAGGAGGAGCATC 281
Db      117  SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY      282  ATGAAGCTGAATGCCACCGAGTTCGAGAACCATGCTGAGGCTCTTACATCCCGCAT 341
Db      137  AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY      342  GAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCGCGAGGG-----GGCTTTGGCTCT 392
Db      157  GluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGln 176
QY      393  CGGGGTACCGCCCGCAGGCTCACCTGTGGACGGGGGCCGCCAGCAGGAGGAGCAAA 452
Db      177  ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY      453  GTGGACATCCCTTCGCTCTGGTGGCCACCGAGTATGTGGTGGCTGCCATTATTGGCAAG 512
Db      194  CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY      513  GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAAAGATAGACGTGCATAGG 572
Db      214  GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY      573  AAGGAGACGCGAGGTGCACCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGC 632
Db      234  LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY      633  TCCTCCGCTGTGAAGATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACACCAACG 692
Db      254  SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
```

QY 513 GAGGGGCCCATCCGCAACATCACAAAAAGAGCCAGTCCAAAGATAGACGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrIleGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGAGAACGAGTGGAGCTGAAAGAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGC 632
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCCTGTTAGATCATCTCGAGATTATGATATAAGAGGCTTAAGGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTCAGCAGGTTCCTCCCTGAGATCCTGGCCCATATAACTTTGTAGGGCGTCTCATTCGC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGAGAGACGGAACCTGAAGAGTAGAGCAGGATACCGAGACAAAAATCACCCTCC 812
 Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 QY 813 TCGTGTGACAGCCTTACCTTTACAACTGAGAGGACCATCAGTGTGAAGGGGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAATTGTCGAGGGCCGAGCAGGAAATATGAAAGATTTCGGGAGCGCCTATCAGAA 932
 Db 334 GluThrCysAlaLysAlaGluGluIleMetLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGGTGCATGAGC-----TCTCACTGTATCCCTGGCTGAACTGCTGCTGTA 986
 Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
 QY 987 GGTCTTTTCCGAGCTTCATCAGCGCAGTCCCGCG-----CCTCCAGCAGCGT 1037
 Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGGTGTCTCCCTATAGCTCTTTTATGACAGGCTCCGAGCAGGAGTGTGTGAGTG 1097
 Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY 1098 TTTATCCCGCCGAGCAGTGGCGCATCATCGGCAAGAGGGGAGCAGCATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyIleHisIleLysGln 430
 QY 1158 CTCTCCCGTTTCCGAGCGCTCCATCAAGATTGCACCCGCAACACCTCACTCCAAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCGTATGTTTATCATCTGACCGCCGAGGCGCCCAATTCAGGCTCAGGGAAGAATC 1277
 Db 451 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGAACACTCAAGAGAGAACTTCTTTGTTGCCAAGGAGGAAGTGAAGCTGAGAC 1337
 Db 471 TyrGlyLysIleLysGluLeuAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY 1338 CACATACGTGTGCACATCAGCAGCTGCGCGGTCTATGGCAAGAGTGGAAAAAGGTG 1397
 Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleLysGlyGlyThrVal 510
 QY 1398 AACGAGTTGCAGATTTCAGCGAGCTGAGGTGTAGTAGTACCAAGAGACAGACCCCTGAT 1457
 Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY 1458 GAGAACGACAGGTCTCGTGAATAATCATCGGCATTTCTATCCAGTCAGATGCTCAA 1517
 Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY 1518 CGAAGATCCGACATCTCTGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
 Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570

QY 1572 AGTAACACGAGCCGAGGACGAGGAG 1598
 Db 571 SerGlyProGlnSerArgLys 579
 RESULT 5
 US-09-643-597-176
 ; Sequence 176, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; CURRENT FILING DATE: 2000-08-21
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 176
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-643-597-176
 Alignment Scores:
 Pred. No.: 8,4e-161 Length: 579
 Score: 1673.00 Matches: 380
 Percent Similarity: 84.48% Conservative: 50
 Best local Similarity: 74.66% Mismatches: 63
 Query Match: 53.04% Indels: 16
 DB: 4 Gaps: 9
 US-09-270-437D-7 (1-1946) x US-09-643-597-176 (1-579)
 QY 102 CGGTTAGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATCGGAGTA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnIleVal 96
 QY 162 CTGGACAGCTCTGCTGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
 QY 222 AGTGAGACGCGAGTGTGTAATGTCCCTATTTCCACCGGGAGCAGACCGAGGAGCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACAGTGGAGAACCATGCCCTCAAGTCTCTACATCCCGAT 341
 Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaIlePheProAsp 156
 QY 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCGCAGGG-----GCCTTTGGCTCT 392
 Db 157 GluMetAlaAlaGlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGln 176
 QY 393 CGGGGTACGCGCGCGAGGCTCACCTGTGGCAGCGGGGGCCCGCCAGCAGGAGCAAC 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGGACATCCCTCTCGCTCTGTCGCCACAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 512
 Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleLeuGlyLys 213
 QY 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572

```
Db 177 ArgGlySerSerArgGlnGlnSerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCCGCTCTGGTCCACCCAGCATGTGGTGCCTATTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleGlyLys 213
QY 513 GAGGGGCCCATCCCGCAATCATCAAAACAGACCAGTCCCAAGATAGAGCGTCATAG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACGCGAGTGCAGTGCAGAAAGCCATCAGTGCACCTCCACCCCTGAGGGCTGC 632
Db 234 LysGluAsnAlaGlyAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTGTGAAGATCTTGGAGATTATGCATAAAGAGCTTAAGGACACCAAAACG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGTCCCTGAGATCCTGGCCCATAACTTTGTAGGCGCTCTCATGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAGGAGCGAAGCTGGAAGAGGTAGACAGATACCGAGACAAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCCTTGACAGCTTACCTTTTACACCTTGAGAGGACCATCTGTGAAGGGGCGCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATGTTCGAGGCGGACGAGGAATAATCAGAAAGTTCGGGAGGCGCTATGAAAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGCTGCCATGAGC-----TCTCACCTGATCCTGGCTGCACTGGCTGTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGGCGAGTCCGCGCG-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGCGCTCCTCCATAGCTCCTTTATCAGGCTCCCGAGGAGATGTCGAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGGAGTGGCGCCATCATCGGCAAGAGGGGAGGAGCAGCATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln 430
QY 1158 CTCTCCCGTTTCCAGCGCTCTCATCAAGATTGCAACCCCGCAAAACACTGACTCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCACTGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAAACTCAAGGAGGAGAACTTCTTTGTGCCCAAGGAGGAAGTGAAGCTGGAGAC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATAGTGTGCAGCATCAGCGCTGGCGGGTCTATTGGCAAAAGTGAAGAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAGATTTCAGCGAGCTGAGTGGTAGTACCAAGAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACGAGTTCATCGTGAATAATCATCCGACATTTCTATGCCAGTCAGATGGCTCAA 1517
Db 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
```

```
QY 1518 CGGAAGATCCGAGACATCTCGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACGAGCCCGCAGCAGCAGGAGAG 1598
Db 571 SerGlyProGlnSerArgArgLys 579
RESULT 4
US-09-606-421B-348
; Sequence 348, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348
Alignment Scores:
Pred. No.: 5,54e-161 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 4 Gaps: 9
US-09-270-437D-7 (1-1946) x US-09-606-421B-348 (1-579)
QY 102 CGGGTTAGAGCGCGGAAATTCAAATCCGAAATATTCACCCCGAGCTCCGATCGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGGACAGCTGCTGCTCAGTATGTA CAGTAGAAGTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCATGTGTGATGTCACCTATTCCAACCGGAGCAGACCGAGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaLeu 136
QY 282 ATGAAGCTCAATGGCCACCGAGTTGGAGAACCATGCCCTGAAGTCTCTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGGCGCGGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGTGCACCCCGCCAGGCGCTCAGCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCCGCTCTGGTCCACCCAGCATGTGGTGCCTATTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuValProThrGlnPheValGlyAlaIleGlyLys 213
```

Db 137 AsplysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA--CAGGACCTCTGAGAAATGGCGCCGAGGG-----GGCTTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGln 176
 QY 393 CGGGGTGAGCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCGAGCCAGCAGAGCAA 452
 Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
 QY 453 GTGACATCCCTCCCTCGCTCTGGTGGCCAGCCAGTATGTGGTGGCCATTATTGGCAAG 512
 Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
 QY 513 GAGGGGCCACATCCGCAACATCACAAACAGACCCAGTCCAGATAGAGTGCATAGG 572
 Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
 QY 573 AAGGAGAACGAGTGCAGCTGAAAGAACCCATGAGTGTGCACTCCACCCCTGAGGGCTGC 632
 Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
 QY 633 TCCTCCGCTTTAAGATCATCTGGAGATTATGCATTAAGAGGCTTAAGGACACCAAAACG 692
 Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
 QY 693 GCTGACGAGGTTCCCTCAAGATCTCGCCCATATAACTTTGTAGGGCGTCTCATTTGCC 752
 Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
 QY 753 AAGGAGAACGGAACCTGAAGAGGTAGAGAAGATACCGAGACAAATAATCATCTCTCC 812
 Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
 QY 813 TCGTTGACAGCTTACCTTTACACCTGAGAGGACCATCAGTGTGAAGGGGCCATC 872
 Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
 QY 873 GAGAAATCTGTGAGGGCCGAGCAGGAAATATGAAGAAGTTTCGGGAGGCTTATGAGAA 932
 Db 334 GluThrCysAlaLysAlaGluGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
 QY 933 GATGTGGTGCATGAGC-----TCTCACTGTATCCCTGGCTGAACTGCTGCTGTA 986
 Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 373
 QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGCGTT 1037
 Db 374 GlyLeuPheProProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
 QY 1038 ACTGGGGTGTCTCCCTATAGCTCTTTTATGCAAGGCTCCGAGCAGGATGCTGAGGTG 1097
 Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
 QY 1098 TTTATCCCGCCAGGAGTGGGCGCATCATCGGAGAGAGGGGAGCAGACATCAACAG 1157
 Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGln 430
 QY 1158 CTCTCCCGTTTGGCAGCGCTTCATCAGATTGGACACCCGAAACCTGACTCCAA 1217
 Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
 QY 1218 GTTCTATGTTTATCATCTGAGCGCCGAGAGGGCCCAATTCAGAGCTCAGGAGAAATC 1277
 Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
 QY 1278 TATGGCAACTCAAGAGAGAACTTCTTTGGTCCCAAGAGAGAGTGAAGCTGGAGACC 1337
 Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
 QY 1338 CACATAGCTGTCCAGCATCAGCAGTGGCGGGTCAATTGGCAAGAGTGGAAACCGGTG 1397
 Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510

QY 1398 AACAGATTGCAGATTTCACCGCAGCTGAGTGGTAGTACCAAGAGACCCAGACCCCTGAT 1457
 Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
 QY 1458 GAGAACACACAGGTTCATCGTGAATAATCATCGACATTTCTATGTCAGTTCAGATGGCTCAA 1517
 Db 531 GluAsnAspGlnValValValValIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
 QY 1518 CGAAGATCCGAGACATCTCGCCCGAGGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
 Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
 QY 1572 AGTAACACGCGCCAGGACGAGGAG 1598
 Db 571 SerGlyProProGlnSerArgArgLys 579
 RESULT 3
 ; Sequence 348, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121-455C8
 ; CURRENT APPLICATION NUMBER: US/09/542,615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 348
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-542-615A-348
 Alignment Scores:
 Pred. No.: 5,54e-161 Length: 579
 Score: 1875.00 Matches: 380
 Percent Similarity: 84.68% Conservative: 51
 Best Local Similarity: 74.66% Mismatches: 62
 Query Match: 53.10% Indels: 16
 DB: 4 Gaps: 9
 US-09-270-437D-7 (1-1946) x US-09-542-615A-348 (1-579)
 QY 102 CGGGTTAGGAGCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTA 161
 Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
 QY 162 CTGGACAGCTCTCGCTCAGTATGTCAGTACAGTGTGAGCACTGTGAGCACTGACACCGAG 221
 Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGlnGlnValAsnThrAsp 116
 QY 222 AGTGAGCGGAGTGTGATCTCACTTCCACCGGAGCAGACCCAGGCAAGCCATC 281
 Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
 QY 282 ATGAAGCTGAATGGCCACCCAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGAT 341
 Db 137 AsplysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
 QY 342 GAGCAGATAGCA--CAGGACCTGAGATGGCGCCGAGGG-----GGCTTTGGCTCT 392
 Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
 QY 393 CGGGGTGAGCCCGCCGAGGGCTCACCTGTGGCAGCGGGGGCCCGAGCAGAGCAA 452

QY 228 ACGCAGTGGTGAATGTCACCTATTCCAAACCGGAGCAGACCGAGCAAGCCATCATGAAG 287
Db |||||
119 ThrAlaValValAsnValThrTy-SerAsnArgGluGlnThrArgGlnAlaIleMetLys 138
QY 288 CTGAATGGCCACCATGTTGGAGAACCATGCCCTGAAGTCTCTCATATCCCCCGATGAGCAG 347
Db |||||
139 LeuAsnGlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyriIleProAspGluGln 158
QY 348 ATAGCACAGGACCTGAGAAATGGCGCGCGAGGGGGCTTTGGCTCTCCGGGGFCACCCCGC 407
Db |||||
159 IleThrGlnGlyProGluAsnGlyArgArgGlyPheGlySerArgGlyGlnProArg 178
QY 408 CAGGGCTACCTGTGGCAGCGGGGGCCCGCCAGCAGCAGCAAGTGGACATCCCCCTT 467
Db |||||
179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 468 CGGCTCTGGTGGCCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGGCCACCATC 527
Db |||||
199 ArgLeuLeuValProThrGlnTyriValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 528 CGCAACATCAAAACAGACCCAGCTCCAAGATAGACGTGCATAGAGAGAGAAACCGCAGGT 587
Db |||||
219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 588 GCAGCTGAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTCTCTCCCGTTGTAAG 647
Db |||||
239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 648 ATGATCTTGGAGATTATGCATAAAGAGCTAAGACACCAAAACGGCTGACGAGGTCCC 707
Db |||||
259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 708 CTGAAGATCTCGGCCATAATACTTTGTAGGGCGTCTCATTTGGCAAGAGAGACGGAAC 767
Db |||||
279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 768 CTGAAGAGGTAGACAGATACCGACACAAAATCACCATCTCTCTGTTGCAAGACCTT 827
Db |||||
299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 828 ACCCTTTACACCTGAGAGACCATCAGTGTGAAGGGGGCCATCGAGAATTTGTCAGG 887
Db |||||
319 ThrLeuTyranProGluuArgThrIleThrValLysGlyAlaIleGluAsnLysCysArg 338
QY 888 GCGCAGCAGGAAATATCAAGAAAGTTCCGAGGCGCTTATGAGATGATGTGGCTGCCATG 947
Db |||||
339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyriGluAsnAspValAlaAlaMet 358
QY 948 AGC-----TCTACCTGATCCCTGGCGTGAACCTGGCTGTGTAGTCTTTTCCAGCT 1001
Db |||||
359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValAlaGlyLeuPheProAla 378
QY 1002 TCATCCAGCGCAGTCCCGCGCTCCAGCAGCGCTTACTGGGGCTGTCTCCCTATAGCTCC 1061
Db |||||
379 SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTyriSerSer 398
QY 1062 TTATGACAGGCTCCGAGCAGGAGATGTGCAGGTGTTATCCCGCGCCACGAGCAGTGGGC 1121
Db |||||
399 PheMetGlnAlaProGluGlnGluMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1122 GCCATCATCCGCAGAGGGGCAGCACATCAACAGCTCTCCCGGTTTGGCAGCGCTCC 1181
Db |||||
419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1182 ATCAAGATTGCACACCGCAACACCTGACTCCAAAGTTCGTATGTTTATCATCTGGA 1241
Db |||||
439 IleLysIleAlaProProGluThrProAspSerLysValArgMetValIleThrGly 458
QY 1242 CCGCCAGAGGCCCAATTCAGGGCTCAGGGAAGAATCTATGCGCAAACTCAAGGAGGAGAAC 1301
Db |||||
459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyriGlyLysLeuLysGluGluAsn 478

QY 1302 TTCTTTGGTCCCAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
Db |||||
479 PhePheGlyProLysGluGluValLysLeuGluThrHisIleArgValProAlaSerAla 498
QY 1362 GCTGGCCGGGTTCATTGGCAAGGTGGAAAAACGGTGAACGAGTTGCAGAAATTTGACCGCA 1421
Db |||||
499 AlaGlyArgValIleGlyLysGlyGlyThrValAsnGluLeuGlnAsnLeuThrAla 518
QY 1422 GCTGAGGTGTAGTACCACAGAGACACAGCCCTGTAGAGAACACACAGGTTCATCGTCAAA 1481
Db |||||
519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValIleValLys 538
QY 1482 ATCATCGGACATTCTTATGCGAGTCCAGATCGGTCAACGGAAGATCCGAGACATCTCGCC 1541
Db |||||
539 IleIleGlyHisPheTyriAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
QY 1542 CAGGTTAAAGCAGCAGCATCAGAGGAGGACAGAGTAACAGGCCCGCAGGACCGAGGAAG 1598
Db |||||
559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgLys 577

RESULT 2
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.: 5,54e-161 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 4 Gaps: 9

US-09-270-437D-7 (1-1946) x US-09-643-597-348 (1-579)

QY 102 CGGGTTAGGAGCCGGAAATTCGAATCCGAATATTCACCCAGCTCCGATGGAGTA 161
Db |||||
77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluVal 96
QY 162 CTGCAGACGCTGTGGCTCAGTATGTGTACAGTAGAAGTGTGAGCAAGTGAACACCGAG 221
Db |||||
97 LeuAspSerLeuLeuValGlnTyriGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGTGATGTCCATTCATTCACCGGAGCAGACCGAGCAAGCATC 281
Db |||||
117 SerGluThrAlaValValAsnValThrTyriSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCGCCACCGTTGGAGAACCATGCCCTGAGGTCTTCATCATCCCGAT 341
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:56:36 ; Search time 17.4913 Seconds
(without alignments)
11487.328 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctagaggaggctggg.....attccttcagggttttaaaa 1946

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0270437/runat_23072004_095740_27466/app_query.fasta_1.5582
-DB=Issued Patents AA -QFMT=fastan -SURFIX=ral -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0270437@cgn 1.1 59 @runat_23072004_095740_27466 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/ECTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2491	70.5	577	3	US-09-261-855-2
2	1875	53.1	579	4	US-09-643-597-348
3	1875	53.1	579	4	US-09-542-615A-348
4	1875	53.1	579	4	US-09-606-421B-348
5	1873	53.0	579	4	US-09-643-597-176
6	1873	53.0	579	4	US-09-480-884A-176
7	1873	53.0	579	4	US-09-542-615A-176
8	1873	53.0	579	4	US-09-606-421B-176
9	248.5	7.0	643	1	US-08-021-608D-10
10	248.5	7.0	643	1	US-08-726-160-10
11	248.5	7.0	643	5	PCT-US94-01782-10
12	245	6.9	49	3	US-09-261-855-18

13	245	6.9	644	1	US-08-021-608D-2	Sequence 2, Appli
14	245	6.9	644	1	US-08-726-160-2	Sequence 2, Appli
15	245	6.9	644	5	PCT-US94-01782-2	Sequence 2, Appli
16	242	6.9	590	1	US-08-021-608D-8	Sequence 8, Appli
17	242	6.9	590	1	US-08-726-160-8	Sequence 8, Appli
18	242	6.9	590	5	PCT-US94-01782-8	Sequence 8, Appli
19	236	6.7	48	3	US-09-261-855-20	Sequence 20, Appli
20	229	6.5	47	3	US-09-261-855-17	Sequence 17, Appli
21	228	6.5	47	3	US-09-261-855-19	Sequence 19, Appli
22	222	6.3	49	3	US-09-261-855-22	Sequence 22, Appli
23	221.5	6.3	530	1	US-08-187-793-4	Sequence 4, Appli
24	219	6.2	48	3	US-09-261-855-24	Sequence 24, Appli
25	214	6.1	47	3	US-09-261-855-21	Sequence 21, Appli
26	191	5.4	47	3	US-09-261-855-23	Sequence 23, Appli
27	174.5	4.9	1418	4	US-09-252-991A-32367	Sequence 32367, A
28	174	4.9	343	1	US-08-187-793-2	Sequence 2, Appli
29	169.5	4.8	577	4	US-09-252-991A-19178	Sequence 19178, A
30	169	4.7	720	4	US-09-252-991A-21881	Sequence 21881, A
31	165	4.7	657	4	US-09-252-991A-28001	Sequence 28001, A
32	162.5	4.6	489	4	US-09-252-991A-16940	Sequence 16940, A
33	162.5	4.6	697	4	US-09-252-991A-24009	Sequence 24009, A
34	160	4.5	774	4	US-09-252-991A-16789	Sequence 16789, A
35	159	4.5	1008	4	US-09-252-991A-29419	Sequence 29419, A
36	157	4.4	782	4	US-09-252-991A-30464	Sequence 30464, A
37	154.5	4.4	243	1	US-08-021-608D-6	Sequence 6, Appli
38	154.5	4.4	243	1	US-08-726-160-6	Sequence 6, Appli
39	154.5	4.4	243	5	PCT-US94-01782-6	Sequence 6, Appli
40	154.5	4.4	904	4	US-09-976-594-615	Sequence 615, App
41	153	4.3	1093	4	US-09-252-991A-21827	Sequence 21827, A
42	151	4.3	1411	4	US-09-252-991A-28408	Sequence 28408, A
43	150.5	4.3	957	4	US-09-252-991A-20408	Sequence 20408, A
44	149	4.2	1706	4	US-09-252-991A-31760	Sequence 31760, A
45	148.5	4.2	563	4	US-09-252-991A-28458	Sequence 28458, A

ALIGNMENTS

RESULT 1

US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:
Pred. No.: 1.01e-216 Length: 577
Score: 2491.00 Matches: 493
Percent Similarity: 99.00% Conservative: 1
Best Local Similarity: 98.80% Mismatches: 3
Query Match: 70.55% Indels: 2
DB: 3 Gaps: 1

US-09-270-437D-7 (1-1946) x US-09-261-855-2 (1-577)

Qy	108	AGGAGCGGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAGTACTGGAC	167
Db	79	ArgSerArgLysileGlnileArgAsnilleProGlnleuArgirpGluValleuasp	98
Qy	168	AGCTCTGCTGCTCAGTATGTGACAGTAGAGAACTGTGAGCAAGTGAACACCGAGGTGAG	227
Db	99	SerleuLeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGlu	118

This Page Blank (uspto)


```
QY 933 GATGTGGCTCCCATGAGC-----TCTCACTGATCCCTGGCCTGAACTGGCTGTGTA 986
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 361 AspileAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 380
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCG-----CCTCCAGCAGGTT 1037
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 381 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 400
QY 1038 ACTGGGCTGCTCCCTATAGCTCTCTTATCAGCTCCCGCAGCAGAGATGGTGCAGGTG 1097
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 401 Thr-----ProProlyrProGlnPheGluGluSer---GluThrGluThrValHisLeu 417
QY 1098 TTTATCCCGCCCGCAGCGAGTGGCGCCATCATCGCAAGAAAGGGCGCAGCATCAACAG 1157
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 418 PheileProAlaLeuSerValGlyAlaIleGlyLeuGlnGlyHisIleLysGln 437
QY 1158 CTCTCCGGTTTCCAGCGCTTCATCAAGATTGCAACCCCGAACAACCTGACTCCAA 1217
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 438 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 457
QY 1218 GTTCGTATGTTATCATCTGACCGCCAGCGCCCAATTCAAGGCTCAGGGAAGAATC 1277
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 458 ValArgMetValIlelleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 477
QY 1278 TATGGCAAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAAC 1337
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 478 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 497
QY 1338 CATATAGTGTGTCAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGAAGAGGTG 1397
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 498 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 517
QY 1398 AACAGAGTGCAGATTTCAGCGCAGCTGAGGTGTAGTACCAAGACACAGACCCCTGAT 1457
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 518 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 537
QY 1458 GAGAACGACAGGTATCTGTAATAATCATCGCAATTTCTATGCTCAGTCAAGTGGCTCAA 1517
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 538 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 557
QY 1518 CGNAGATCCGAGACATCTCGGCCCGGTTAAGCAG---CAGCATCAGAAAGGA---CAG 1571
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 558 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 577
QY 1572 AGTAACAGCGCCAGGACGAGGAAG 1598
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 578 SerGlyProProGlnSerArgArgLys 586

RESULT 15
ID AAB11328
XX AC AAB11328;
XX AC AAB11328;
XX AC AAB11328;
DT 21-FEB-2001 (first entry)
XX DE Human lung cancer-associated protein L523S.
XX DE Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
XX DE vaccine; detection.
XX OS Homo sapiens.
XX PN WO200061612-A2.
XX PD 19-OCT-2000.
XX PF 03-APR-2000; 2000WO-US008896.
XX PR 02-APR-1999; 99US-00285479.
XX PR 17-DEC-1999; 99US-00466396.
XX PR 30-DEC-1999; 99US-00476496.
```

```
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PA (CORI-) CORIXA CORP.
XX Wang T, Fan L;
XX WPI; 2000-628399/60.
DR N-PSDB; AAC65900.
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT protein is used for detecting and monitoring progression of lung cancer
PT in a patient.
XX Claim 3; Page 186-188; 261pp; English.
XX This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2, and then administered to the patient to inhibit
CC development of cancer
XX Sequence 579 AA;

Alignment Scores:
Pred. No.: 9,35e-160 Length: 579
Score: 1873.00 Matches: 380
Percent Similarity: 84.48% Conservative: 50
Best Local Similarity: 74.66% Mismatches: 63
Query Match: 53.04% Indels: 16
DB: Gaps: 9

US-09-270-437D-7 (1-1946) x AAB11328 (1-579)
```

```
QY 102 CGGGTATTAGGCGCGAAATTCAAATCCGAAATATTCACCCCGCTCCGATGGGAAGTA 161
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnThrGluVal 96
QY 162 CTGGACAGCTGCTGGCTCAGTATGTACAGAACTGTGAGAACTGTGAGCAAGTGAACACGAG 221
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGACAGCGCAGTGTGATGTCACTTATCCAAACGGGAGCAGACCCAGGCAAGCATC 281
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGTGAATGGCCACCATGTTGGAGAACCATGCCCTGAAGTCTCTACATCCCCGAT 341
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGACCTGAGAAATGGGCGCGAGGG-----GGCTTTGGCTCT 392
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 157 GluMetAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGln 176
QY 393 CGGGGTTCAGCCCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCCAGCAGCAGCA 452
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGGCTCTCGGTGCCACCCAGTATGTGGTGCCATTATTGGCAAG 512
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIlelleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
DB |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGAACCGCAGGTGCAGCTGAAAGGCCATCAGTGTGCATCCACCCCTGAGGGCTGC 632
```

```

QY      1458  GAGAAACGACGAGTTCATCGTGAAAATCATCGACATTCTATGCCAGTCAGATGGCTCAA 1517
Db      538   GluAsnAspGlnValValValLysIleThrGlyHisPheTyxAlaCysGlnValAlaGln 557
QY      1518  CGGAAGATCCGAGACATCTCTGGCCCCAGGTTAAGCAG---CAGCATCAAGAAGGGA---CAG 1571
Db      558   ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 577
QY      1572  AGTAAACCAGCCCGCAGCACCGACCGAGNAG 1598
Db      578   SerGlyProProGlnSerArgArgLys 586

RESULT 14
ADAX28517
ID    ADA28517 standard; protein; 586 AA.
XX
XX    ADA28517;
XX
XX    20-NOV-2003 (first entry)
XX
DE    Recombinant human lung tumour protein L523S.
XX
XX    cancer; lung cancer; gene therapy; vaccine; human;
KW    lung squamous cell carcinoma.
XX
OS    Homo sapiens.
XX
XX    US2003064947-A1.
XX
XX    03-APR-2003.
XX
XX    30-NOV-2001; 2001US-00007700.
XX
XX    18-MAR-1998; 98US-00040802.
XX    27-JUL-1998; 98US-00123912.
XX    22-DEC-1998; 98US-00221107.
XX    02-APR-1999; 99US-00285479.
XX    17-DEC-1999; 99US-00466396.
XX    30-DEC-1999; 99US-00476496.
XX    10-JAN-2000; 2000US-00480884.
XX    22-FEB-2000; 2000US-00510376.
XX    04-APR-2000; 2000US-00542615.
XX    28-JUN-2000; 2000US-00606421.
XX    02-AUG-2000; 2000US-00630940.
XX    21-AUG-2000; 2000US-00643597.
XX    15-SEP-2000; 2000US-00662786.
XX    09-OCT-2000; 2000US-00685696.
XX    12-DEC-2000; 2000US-00735705.
XX    07-MAY-2001; 2001US-00850716.
XX    28-JUN-2001; 2001US-00897778.
XX
XX    (CORI-) CORIXA CORP.
XX
XX    Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI    Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI    Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
XX    WPI; 2003-540798/51.
XX
XX    New isolated polynucleotides and polypeptides useful for diagnosing,
PT    preventing and/or treating cancer, particularly lung cancer.
XX
XX    Claim 9; Page 273-275; 296pp; English.
XX
XX    The invention describes isolated polynucleotides and polypeptides useful
CC    for diagnosing, preventing and/or treating cancer, particularly lung
CC    cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC    defined nucleotide sequences (e.g. 1012, 900 or 273 bp) given in the
CC    specification; complements of the nucleotide sequences cited above; at
CC    least 10 contiguous residues of the nucleotide sequences under highly
CC    sequence that hybridise to a sequence that is at least 75 or 90% identical to
CC    stringent conditions; a sequence that is at least 75 or 90% identical to

```


DR N-PSDB; ABL49283.
XX polynucleotides encoding lung tumor polypeptides, useful for treating
PT lung cancer or stimulating an immune response.
PS
XX Claim 2; Page 354-355; 374pp; English.
XX
CC The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC proteins, T cell populations, or antigen presenting cells that express
CC the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC ABB75070 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 586 AA;

Alignment Scores:
Pred. No.: 6.2e-160 Length: 586
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 5 Gaps: 9

US-09-270-437D-7 (1-1946) x ABB75048 (1-586)

Qy 102 CGGTTAGAGCGCGGAAATTCMAATCCGAATATTCACCCACCTCCGATGGGAAGTA 161
Db 84 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 103

Qy 162 CTGGACACCTGCTGGCTCAGTAGGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
Db 104 LeuAspSerLeuLeuValGlnTrpGlyValValGluSerCysGluGlnValAsnThrAsp 123

Qy 222 AGTCAGACCGCAGTGGTGAATGTCACTATTCCAAACCGGAGCAGACCCAGCAACCCATC 281
Db 124 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 143

Qy 282 ATGAAGCTGAATGGCCACAGTTGGAGAACCCATGCCCTGAAGTCTCTACATCCCGAT 341
Db 144 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 163

Qy 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCCGAGGG-----GGCTTTGGCTCT 392
Db 164 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 183

Qy 393 CGGGTTCAGCCCGCAGCGGCTCACCTGTGGCAGCGGGGGCCCCAGCCCAAGCAGCAAA 452
Db 184 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 200

Qy 453 GTGGACATCCCTTCGGCTCTGGTGCCACCCAGTAGTGTGGTGGCCATTTATTTGGCAG 512
Db 201 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 220

Qy 513 GAGGGGGCCACCATCCGCAACATCACAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db 221 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 240

Qy 573 AAGGAGACGCGAGGTGCAGCTGAAAGAACCCATCAGTGTGCATCCACCCCTCAGGGCTGC 632
Db 241 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 260

Qy 633 TCCTCGCTGTGAAGATGATCTTGGAGATTATGAGAGTATGAGAGTAAAGACCCCAAACG 692
Db 261 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 280

Qy 693 GCTGACGAGGTTCCTCGAAGTCTCTGGGCCCATATAACTTTCTAGGGGCTCTCAPTGGC 752
Db 281 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 300

Qy 753 AAGGAAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAATAATCACCATCTCC 812
Db 301 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 320

Qy 813 TCGTTTCAAGACCTTACCCCTTTACACCCCTGAGAGGACCATCCTGTGAAGGGGCCATC 872
Db 321 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 340

Qy 873 GAGAAATGTTTCAGGGCGGACGAGCAATAATATGAAAGTTTCGGGAGCCCTATCAGAT 932
Db 341 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 360

Qy 933 GATGTGGTGCACAGC-----TCTCACCTGATCCCTGCGCTGAACCTGCTCTCTGTA 986
Db 361 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeu 380

Qy 987 GGTCTTTTCCAGGCTTCATCCAGCGCAGTCCCGCG-----CCTCCACGACGCTT 1037
Db 381 GlyLeuPheProProThrSerSerGlyMetProProProThrSerGlyProProSerAlaMet 400

Qy 1038 ACTGGGGTGTCTCCTATAGCTCCTTTATGAGGCTCCGAGCAGGAGATGTCGAGGTG 1097
Db 401 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 417

Qy 1098 TTTATCCCGCCAGGAGTGGCGCCATCATCGCAAGAGGGGCGAGCACATCAAAACAG 1157
Db 418 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 437

Qy 1158 CTCTCCCGGTTCGCGAGCGCTCCATCAAGATTGCACCCCGCAAAACCTGACTCCAAA 1217
Db 438 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 457

Qy 1218 GTTCGTATGTTATCATCTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGATC 1277
Db 458 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 477

Qy 1278 TATGCCAACTCAAGGAGGAACCTTCTTGTGTCCAAGGAGAGTGAAGCTGGAGACC 1337
Db 478 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 497

Qy 1338 CACATACGTGTGCCAGCATCAGCAGCTGGCGCGGTCATTGGCAAAAGGTGAAAAACGGTG 1397
Db 498 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 517

Qy 1398 AACGATGTCGAAATTTGACGCGCAGCTGAGTGTGTAGTACCAAGAGACAGACCCCTGAT 1457
Db 518 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 537

Qy 1458 GAGAACGCCAGGTTCATCGTGAATATCATCGGACATTTCTATGCCAGTCAGATGGCTCAA 1517
Db 538 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 557

Qy 1518 CGGAAGATCCGAGACATCTCTGGCCCGAGGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db 558 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 577

Qy 1572 AGTAACGAGGCCAGGACCGAGGAAG 1598
Db 578 SerGlyProProGlnSerArgArgLys 586

RESULT 13
ABP61968
ID ABP61968 standard; protein; 586 AA.
XX
AC ABP61968;
DT
XX 07-OCT-2002 (first entry)
DE Human lung cancer associated protein sequence SEQ ID NO:427.
XX Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
XX Homo sapiens.
OS

Qy	1098	TTTTATCCCGCCAGCAGTGGCGGCATCATCGGCAAGAGGAGGCGACATCAACAG	1157
		: : : : : : : : : : : : :	
Db	411	PheileProAlaLeuSerValGlyAlaIlelleGlyLysGlnGlnHisIleLysGln	430
Qy	1158	CTCTCCCGGTTTGCAGCGGCTCCATCAAGATTCACACCCGAAAAACCTGACTCCAAA	1217
Db	431	LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys	450
Qy	1218	GTTTCGTATGTTATCATCATCTGGACCCAGAGAGGCCCAATTCAAAGGCTCAGGGAAGAATC	1277
Db	451	ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle	470
Qy	1278	TATGGCAAACTCAGGAGGAGCACTCTTTGGTCCCAAGGAGGAGTCAAGCTCGAGACC	1337
Db	471	TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla	490
Qy	1338	CACATACGTGTGCAGCATCAGCAGCTGGCGCGGTCAATTGGCAAGGTGAAAAACGGTG	1397
Db	491	HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal	510
Qy	1398	AACAGTTGCAGAAATTTGACGGCAGCTGAGGTTGTTAGTACACAGAGACCAGACCCCTGAT	1457
Db	511	AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp	530
Qy	1458	GAGAACCAACAGGTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCCAGTGGCTCAA	1517
Db	531	GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln	550
Qy	1518	CGGAAGATCCGAGACATCTCTGGCCCGAGTTAAGCAG---CAGCATCAGAAAGGGA---CAG	1571
Db	551	ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln	570
Qy	1572	AGTAACAGCGCCAGGACCGGACCGAGGAG	1598
Db	571	SerGlyProGlnSerArgArgLys	579
RESULT 12			
ABB75048			
ID	ABB75048 standard; protein; 586 AA.		
XX	ABB75048;		
XX	AC AC		
XX	01-MAY-2002 (first entry)		
DT			
DE	Human lung tumour L523S recombinant protein sequence SEQ ID NO:427.		
XX	Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;		
KW	immune response.		
XX	Homo sapiens.		
OS			
XX	WO200200174-A2.		
PN			
XX	03-JAN-2002.		
PD			
XX	28-JUN-2001; 2001WO-US021065.		
PF			
XX	28-JUN-2000; 2000US-00606421.		
PR	02-AUG-2000; 2000US-00603940.		
PR	21-AUG-2000; 2000US-00643597.		
PR	15-SEP-2000; 2000US-00662786.		
PR	09-OCT-2000; 2000US-00685696.		
PR	12-DEC-2000; 2000US-00735705.		
PR	07-MAY-2001; 2001US-00850716.		
XX	(CORI-) CORIXA CORP.		
PA			
XX	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;		
PI	McNeill PD, Fanger N, Retter MW, Warnerakis M, Fanger GR;		
PI	Vedvick TS, Carter D, Watanabe Y, Peckham DW;		
XX	WPI; 2002-090513/12.		
DR			

Db	234	LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr	253
Qy	633	TCCTCCGCTTGAAGATGATCTTGGAGATTATGATTAAGAGGCTTAAGGACACCAAAACG	692
Db	254	SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe	273
Qy	693	GCTGACGAGTTCCTCCTGAAGATCCTGCCCATATAACTTTGTAGGCGCTCATTCGCG	752
Db	274	ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly	293
Qy	753	AAGAAGACGACGACCTGAAGAGTAGACAGATACCGAGACAAATAATCACCATCTCC	812
Db	294	LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer	313
Qy	813	TCGTTTGCAGACCTTACCTTTTACAACTGAGAGCACCATCTGTGAAGGGGCCATC	872
Db	314	ProLeuGlnGluLeuThrLeuTyraAsnProGluArgThrIleThrValLysGlyAsnVal	333
Qy	873	GAGAAATGTTGACGGCCGAGCAGGAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAA	932
Db	334	GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn	353
Qy	933	GATGGGTGCATGAGC-----TCTCACTGATCCCTGGCCCTCAACCTGCTCTCTCTA	986
Db	354	AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu	373
Qy	987	GGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCCG-----CCTCCAGCAGCGCTT	1037
Db	374	GlyLeuPheProProThrSerGlyMetProProProProProProProProProSerAlaMet	393
Qy	1038	ACTGGGGTGTCTCCCTATGCTTATGAGCTCCCGAGCTCCGAGCAGGATGTGCGAGTG	1097
Db	394	Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisLeu	410
Qy	1098	TTTATCCCGCCGAGGAGTGGCGCCATCATCGCAAGAGGGCGCACATCAACAG	1157
Db	411	PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGln	430
Qy	1158	CTCTCCCGGTTTGGCAGCGCTCCATCAAGATTGCACACCCGCAACACCTGACTCCAAA	1217
Db	431	LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys	450
Qy	1218	GTTCTGATGGTTATCATCTGACCGCCGAGGCGCCCAATCAAGGCTCAGGAGAAATC	1277
Db	451	ValArgMetValIleIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle	470
Qy	1278	TATGGCAAACTCAAGAGGAGAACTCTTGTGTCACAGGAGGAGTGAAGCTGGAGACC	1337
Db	471	TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla	490
Qy	1338	CACATACGTTGCCAGCATCAGCTGCGCGGCTATTGGCAAAAGGTGGAAAAACGGTG	1397
Db	491	HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal	510
Qy	1398	AACGAGTTGAGAAATTTGACGAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGAT	1457
Db	511	AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp	530
Qy	1458	GAGACGACGAGTCTCTGTAATCATCGACATTTCTATCCAGTCAGATGCTCAA	1517
Db	531	GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln	550
Qy	1518	CGGAAGATCCGAGACATCTCGGCCAGGTGAAGCAG---CAGCATCAGAAAGGGA---	1571
Db	551	ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln	570
Qy	1572	AGTAACGAGCCGACGACGAGGAG	1598
Db	571	SerGlyProProGlnSerArgArgLys	579

RESULT 11
ADA28438

ID	ADA28438	standard; protein; 579 AA.
AC	ADA28438;	
DT	20-NOV-2003	(first entry)
XX	Human lung tumour protein L523S.	
DE	cancer; lung cancer; gene therapy; vaccine; human; lung squamous cell carcinoma.	
KW	Homo sapiens.	
XX	US2003064947-A1.	
PD	03-APR-2003.	
PF	30-NOV-2001; 2001US-00007700.	
XX	18-MAR-1998; 98US-00040802.	
PR	27-JUL-1998; 98US-00123912.	
PR	22-DEC-1998; 98US-00221107.	
PR	02-APR-1999; 99US-00285479.	
PR	17-DEC-1999; 99US-00466396.	
PR	30-DEC-1999; 99US-00476496.	
PR	10-JAN-2000; 2000US-00480884.	
PR	22-FEB-2000; 2000US-00510376.	
PR	04-APR-2000; 2000US-00542615.	
PR	28-JUN-2000; 2000US-00606421.	
PR	02-AUG-2000; 2000US-00630940.	
PR	21-AUG-2000; 2000US-00643597.	
PR	15-SEP-2000; 2000US-00662786.	
PR	09-OCT-2000; 2000US-00685696.	
PR	12-DEC-2000; 2000US-00735705.	
PR	07-JAN-2001; 2001US-00850716.	
PR	28-MAY-2001; 2001US-00897778.	
XX	(CORI-) CORIXA CORP.	
PI	Wang T, Wang A, Skeiky YAM, Li SX, Kalos MD, Henderson RA; McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS; Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;	
XX	WPI; 2003-540798/51.	
DR	N-PSDB; ADA28437.	
XX	New isolated polynucleotides and polypeptides useful for diagnosing, preventing and/or treating cancer, particularly lung cancer.	
PT	Example 2; Page 249-251; 296pp; English.	
XX	The invention describes isolated polynucleotides and polypeptides useful for diagnosing, preventing and/or treating cancer, particularly lung cancer. A new isolated polynucleotide comprises: any of the 22 fully defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the specification; complements of the nucleotide sequences cited above; at least 10 contiguous residues of the nucleotide sequences cited above; a sequence that hybridise to any of the nucleotide sequences under highly stringent conditions; a sequence that is at least 75 or 90% identical to the above nucleotide sequences; or degenerate variants of the above nucleotide sequences. The composition and methods are useful in diagnosing, preventing and/or treating cancer, particularly lung cancer, in gene therapy and in vaccines. This is the amino acid sequence encoded by a human lung tumour cDNA isolated from a lung squamous cell carcinoma that may be useful in the diagnosis and treatment of lung cancer and other disorders.	
SQ	Sequence 579 AA;	
Alignment Scores:	6.17e-160	Length: 579
Pred. No.:	1875.00	Matches: 380
Score:	84.68%	Conservative: 51

```
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
Qy 1158 CTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCAAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
Qy 1218 GTTCGTATGTTATCATCTGAGCCGCGCAGAGCCCAATTCAAAGGTCACGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
Qy 1278 TATGCGAACTCAAGGAGGAGAACTCTTTGGTCCCAAGGAGGAGTCAAGCTCGAGACC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
Qy 1338 CACATAGCTGTCCAGCATCAGCAGCTGGCGGCTCAATTGGCAAGGTGGAACCGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
Qy 1398 AACGAGTTGCAGATTTCAGCGCAGCTGAGTGGTAGTACCAAGAGACCGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
Qy 1458 GAGAACACGAGGTCATCGTGAAATCATCGACATTTCTATGCTCAGTCCAGTGGCTCAA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheYrAlaCysGlnValAlaGln 550
Qy 1518 CGGAAGATCCGAGACATCTCTGGCCAGGTTAAGCAG---CAGCATCAGAAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
Qy 1572 AGTAACGAGCCGACGAGGAGGAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
RESULT 10
ADA28539
ID ADA28539 standard; protein; 579 AA.
AC ADA28539;
XX
XX 20-NOV-2003 (first entry)
XX Recombinant human lung tumour protein L523S #2.
XX cancer; lung cancer; gene therapy; vaccine; human;
XX lung squamous cell carcinoma.
XX Homo sapiens.
XX
XX PN US2003064947-A1.
XX PD 03-APR-2003.
XX PF 30-NOV-2001; 2001US-00007700.
XX PR 18-MAR-1998; 98US-00040802.
XX PR 27-JUL-1998; 98US-00123912.
XX PR 22-DEC-1998; 98US-00221107.
XX PR 02-APR-1999; 99US-00285479.
XX PR 17-DEC-1999; 99US-00466396.
XX PR 30-DEC-1999; 99US-00476486.
XX PR 10-JAN-2000; 2000US-00480884.
XX PR 12-FEB-2000; 2000US-00510376.
XX PR 04-APR-2000; 2000US-00542615.
XX PR 28-JUN-2000; 2000US-00606421.
XX PR 02-AUG-2000; 2000US-00630940.
XX PR 21-AUG-2000; 2000US-00643597.
XX PR 15-SEP-2000; 2000US-00662786.
XX PR 09-OCT-2000; 2000US-00685696.
XX PR 12-DEC-2000; 2000US-00735705.
XX PR 07-MAY-2001; 2001US-00850716.
XX PR 28-JUN-2001; 2001US-00897778.
```

```
XX PA (CORI-) CORIXA CORP.
XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX PI Moneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
XX PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX DR WPI; 2003-540798/51.
XX
XX PT New isolated polynucleotides and polypeptides useful for diagnosing,
XX PT preventing and/or treating cancer, particularly lung cancer.
XX PS Claim 9; Page 288-289; 296pp; English.
XX
XX CC The invention describes isolated polynucleotides and polypeptides useful
XX CC for diagnosing, preventing and/or treating cancer, particularly lung
XX CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
XX CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
XX CC specification; complements of the nucleotide sequences cited above; at
XX CC least 10 contiguous residues of the nucleotide sequences cited above; a
XX CC sequence that hybridise to any of the nucleotide sequences under highly
XX CC stringent conditions; a sequence that is at least 75 or 90% identical to
XX CC the above nucleotide sequences; or degenerate variants of the above
XX CC nucleotide sequences. The composition and methods are useful in
XX CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
XX CC in gene therapy and in vaccines. This is the amino acid sequence of a
XX CC recombinant human lung tumour associated protein.
XX SQ Sequence 579 AA;
Alignment Scores:
Pred. No.: 6.17e-160 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 7 Gaps: 9
US-09-270-437D-7 (1-1946) x ADA28539 (1-579)
```

```
Qy 102 CGGGTAGGAGCGCGAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProHisLeuGlnIleProGluVal 96
Qy 162 CTGGACAGCCTGCTGGCTCAGTATGTGTACAGTACAGAACTGTGAGCAAGTGAACACGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
Qy 222 AGTGACAGCGCAGTGTGATGTCACTTATTCACCGGGAGGAGACGAGGAGCCGATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
Qy 282 ATCAAGCTGAATGGCCACCATGTTGGAGAACCATGCTGAGTCTCTTACATCCCGCAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaIleProAsp 156
Qy 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGGCGCCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGln 176
Qy 393 CGGGGTACGCCCGCCAGGCTCACCTGTGGCAGCGGGGGCCCGCCAGGAGCAGCAGAA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
Qy 453 GTGGACATCCCTCTCGGCTCTCTGGTGGCCACCCAGTATGTGGTGGTCCCATTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
Qy 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
Qy 573 AAGGAGAACGCGAGGTGACGTGAAAGCCATCAGTGTGCATCCACCCCTGAGGCTGC 632
```

Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430

ID	ADA28536
AD	ADA28536 standard; protein; 579 AA.
XX	
AC	ADA28536;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Recombinant human lung tumour protein L523S #1.
XX	
KW	cancer; lung cancer; gene therapy; vaccine; human;
KW	lung squamous cell carcinoma.
OS	Homo sapiens.
XX	
PN	US2003064947-A1.
XX	
PD	03-APR-2003.
PF	
PX	30-NOV-2001; 2001US-00007700.
XX	
PR	18-MAR-1998; 98US-00040802.
PR	27-JUL-1998; 98US-00123912.
PR	22-DEC-1998; 98US-00221107.
PR	02-APR-1999; 99US-00285479.
PR	17-DEC-1999; 99US-00466396.
PR	30-DEC-1999; 99US-00476496.
PR	10-JAN-2000; 2000US-00480884.
PR	22-FEB-2000; 2000US-00510376.
PR	04-APR-2000; 2000US-00542615.
PR	28-JUN-2000; 2000US-00606421.
PR	02-AUG-2000; 2000US-00630940.
PR	21-AUG-2000; 2000US-00643597.
PR	15-SEP-2000; 2000US-00662786.
PR	09-OCT-2000; 2000US-00658696.
PR	12-DEC-2000; 2000US-00735705.
PR	07-MAY-2001; 2001US-00850716.
PR	28-JUN-2001; 2001US-00897778.
XX	(CORI-) CORIXA CORP.
PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI	Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;
PI	Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX	
DR	WPI; 2003-540798/51.
XX	
PT	New isolated polynucleotides and polypeptides useful for diagnosing,
PT	preventing and/or treating cancer, particularly lung cancer.
PS	
PS	Claim 9; Page 285-287; 296pp; English.
XX	
CC	The invention describes isolated polynucleotides and polypeptides useful
CC	for diagnosing, preventing and/or treating cancer, particularly lung
CC	cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC	defined nucleotide sequences (e.g. 1012, 900 or 273 bp) given in the
CC	specification; complements of the nucleotide sequences cited above; at
CC	least 10 contiguous residues of the nucleotide sequences cited above; a
CC	sequence that hybridise to any of the nucleotide sequences under highly
CC	stringent conditions; a sequence that is at least 75 or 90% identical to
CC	the above nucleotide sequences; or degenerate variants of the above
CC	nucleotide sequences. The composition and methods are useful in
CC	diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC	in gene therapy and in vaccines. This is the amino acid sequence of a
CC	recombinant human lung tumour associated protein.
XX	
SQ	Sequence 579 AA;
Alignment Scores:	
Pred. No.:	6,17e-160 Length: 579
Score:	1875.00 Matches: 380
Percent Similarity:	84.68% Conservative: 51
Best Local Similarity:	74.66% Mismatches: 62
Query Match:	53.10% Indels: 16

PI Monell PD, Fanger N, Retter WM, Durham M, Fanger GR, Vedwick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX WPI; 2002-583465/62.
DR N-PSDB; ABQ92483.
XX
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
XX
XX Claim 9; Page 372-374; 381pp; English.

CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
CC invention

XX SQ Sequence 579 AA;

Alignment Scores:

Pred. No.: 6.17e-160 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 5 Gaps: 9

US-09-270-437D-7 (1-1946) x ABP61973 (1-579)

QY 102 CGGTAGGAGCGCGAAATTCGAATTCACCAATTCACCCAGCTCCGATGGGAAGTA 161
DB 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY 162 CTGACAGCCTGCTGGCTAGTATGTGTACGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
DB 97 LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGAGCGAGTGTGAATGTACCTATTCACCGGAGCGAGCAGCGAGCCATC 281
DB 117 SerGluThrAlaValAlaValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAAGCTGAATGCCACCGCTTGGAGAACCATCCCTGAGAGTCTCTACATCCCGAT 341
DB 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCGCAGGG-----GGCTTTGGCTCT 392
DB 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGln 176
QY 393 CGGGGTACGCGCCCGAGGCTACCTGTGCGACCGGGGGCCCCCAGCCAGCAGCAAA 452
DB 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGACATCCCGCTTCCTGGTCCCGACCCAGTATGGGTGCGCATATTGGCAAG 512
DB 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLys 213
QY 513 GAGGGGCGCCACATCCGCAACATCACAACACAGACCCAGTCCAAAGATAGCGTGCATAGG 572
DB 214 GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233

QY 573 AAGAGAACCCAGGTGCAGCTGAAAGAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGC 632
DB 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTTGTAGATGATCTTGGAGATTATCATAAAGAGGCTAAGGACACCAACACG 692
DB 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCTCTGAAGATCTCGGCCCATATAAATCTTTAGGCGCTCTCATGGC 752
DB 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGAGAGCGAACCCTGAAGAGGTAGACAGATACCGAGACAAAAATACCATCTCC 812
DB 294 LysGluGlyArgAsnLeuLysLysIleGluGluAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTGTCAACACCTTACCTTTACACCTGAGAGGACCATCTGTAAGAGGGGCCATC 872
DB 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAATTTGTCAGGGCCGAGCAAGAAATATCAAGAAAGTTCGGGAGGCTATGAGAT 932
DB 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGTGGCTCCCATGAGC-----TCTCACCTGATCCCTGGCTGACCTGGCTGCTGTA 986
DB 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCG-----CTCCACGAGCGGTT 1037
DB 374 GlyLeuPheProThrSerGlyMetProProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGCTCTCCCTATAGCTCTTTATGACGGCTCCCGAGGAGATGGTGAGGTG 1097
DB 394 Thr-----ProProTyrProGlnPheGluInSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGGAGTGGGGCCATCATCGGAAGAGGGGCGAGCATCAACAG 1157
DB 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGGTTTCCAGCGCTCCATCAAGATTGCACCCCGAAACCTCTGACTCCAAA 1217
DB 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCGTATGTTATCATCATCTGACCGCCAGAGCCCAATTCAGGCTCAGGGAAGATC 1277
DB 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGGCAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGAGTGAAGCTGGAGACC 1337
DB 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTGCAGCAGCTGCGCGGTCAATTCGCAAGAGTGAAGGCTGCAAGAGCGTG 1397
DB 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGAGTTGCAATTTGACGCGAGCTGAGGTGGTAGTACCAAGACACAGACCCCTGAT 1457
DB 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACGAGTCTCGTGAATAATCATCGGACATTTCTATCCAGTCCAGATGGCTCAA 1517
DB 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGGAAGATCCGACACATCTCTGGCCCGAGTTAAGCAG---CAGCATCAGAGGGA---CAG 1571
DB 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACAGGCCCGCAGCAGCGAAG 1598
DB 571 SerGlyProGlnSerArgArgLys 579

RESULT 9

CC	ABP61992	represent sequences used in the exemplification of the present
XX	CC	invention
SQ	Sequence 579 AA;	
Alignment Scores:		
Pred. No.:	6.17e-160	Length: 579
Score:	1875.00	Matches: 380
Percent Similarity:	84.68%	Conservative: 51
Best Local Similarity:	74.66%	Mismatches: 62
Query Match:	53.10%	Indels: 16
DB:	5	Gaps: 9
US-09-270-437D-7 (1-1946) x ABP61974 (1-579)		
QY	102	CGGTTTAGGAGCGCGAAATTCMAATCCGAATATTCACCCACGCTCCGATGGGAAGTA 161
DB	77	ArgGlnArgIleAeLgLyLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluVal 96
QY	162	CTGGACAGCCTGTGGCTCAGTATGTGTACGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
DB	97	LeuAspSerLeuLeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY	222	AGTGAGAGCGCAGTGTGAANTGTACCTATTCCAAACCGGAGCAGACACGAGCAAGCCATC 281
DB	117	SerGluThrAlaValAlaValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY	282	ATGAAGCTGTAATGTCACCCAGTGTGAGAACCATGCCCTGAAGTCTCTTACATCCCCGAT 341
DB	137	AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAsp 156
QY	342	GAGCAGATAGCA---CAGGGACCTGAGAAATGGCGCCGAGGG-----GGCTTTGGCTCT 392
DB	157	GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY	393	CGGGTTCAGCCCGCCAGGGCTCAGCTGTGGCAGCGGGGGCCCAAGCAGCAGCA 452
DB	177	ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY	453	GTGGACATCCCGCTCGGCTCTGGTGGCCACCCAGTATGTGGTGCCATTATTGGCAAG 512
DB	194	CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIleLeuGlyLys 213
QY	513	GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCCAAGATACAGCTGATAGG 572
DB	214	GluGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY	573	AAGGAGAACCGAGGTGCAGCTGAAAGACCATCATGTGTGCACTCCACCCCTGAGGGCTGC 632
DB	234	LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY	633	TCTTCCGCTTGAAGATCATCTTGAGATTATGCATTAAGAGGCTTAAGACACCAAAACG 692
DB	254	SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY	693	GCTGACGAGGTTCCCTCAAGATCCTGGGCCATAATACTTTGTAGGGGCTCTCATTTGC 752
DB	274	ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY	753	AAGGAAGGACGGAACCTGAAGAAGGTGAGCAGAGATACCGAGACAAAAATCAACCATCTCC 812
DB	294	LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY	813	TCGTTTGAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCCATC 872
DB	314	ProLeuGlnGluLeuThrLeuTyrAsnProGluAa-gThrIleThrValLysGlyAsnVal 333
QY	873	GAGAAATTGTGCGAGGCGCAGCAGAAATAATGAAGAAAGTTCGGAGGCGCTATGAGAAT 932
DB	334	GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY	933	GATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGGCTGAACCTGGCTGCTGTA 986

Db 254 SerAlaAlaCysLysSerIleLeuLeuIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTCAGCAGGTTCCCTGAAGATCTGGCCCATATAACTTTGTAGGGCGTCTCATTTGGC 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGCAGCACTCTGAAGAGGTAGAGAGATACCGAGACAAATAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGAAGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTTGTCAGGGCCGAGCAGAAATAATGAAGAAAGTTGGGAGGCGCTTATCAGAA 932
Db 334 GluThrCysAlaLysAlaGluGluLeuMetLysLysIleArgGluSerTyArgLys 353
QY 933 GATGTGGTGCATGAGC-----TCTCACTGTATCTCTGGCCTGAACCTGGCTCTCTGA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCTATCCAGCGCAGTCCGCGC-----CCTCCGAGCAGCGTT 1037
Db 374 GlyLeuPheProProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCTATAGCTCTTTATGAGGCTCCGAGCAGCAGATGTCAGGTG 1097
Db 394 Thr-----ProProTyProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGCAGTGGCGCCATCATCGCAAGAAGGGCAGCACATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGln 430
QY 1158 CTCTCCCGTTTGGCAGCGCTCCATCAGATTGCAACCCGCAACACCTGACCTCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTGTATGTTTATCATCTGACCGCCGAGGCGCCCAATTCAGGCTCAGGGAAGATC 1277
Db 451 ValArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470
QY 1278 TATGCAAACTCAAGGAGGAGAACTTCTTTGTCCTCCCAAGGAGAAAGTGAAGCTGAGACC 1337
Db 471 TyrGlyLysIleLysGluAsnPheValSerProLysGluGluValLysLeuGluAla 490
QY 1338 CACATACGTGTCCAGCATCAGCAGCTGGCCGGTTCATTGGCAAGGTGGAAACCGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrVal 510
QY 1398 AACGAGTTGCAGAAATTCAGCGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluValValProArgAspGlnThrProAsp 530
QY 1458 GAGAACGACCAAGGTTCATCGTGAATATCATCGACATTTCTATGCCAGTCAGATGCTCAA 1517
Db 531 GluAsnAspGlnValValLysIleThrGlyHisPheTyAlaCysGlnValAlaGln 550
QY 1518 CGAAGATCCGAGACATCTGTGCCCGCAGGTTAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnLysAlaLeuGln 570
QY 1572 ACTAACCGCGCCGAGCAGGAGGAAG 1598
Db 571 SerGlyProProGlnSerArgArgLys 579
RESULT 6
ABP61917
ID ABP61917 standard; protein; 579 AA.
XX
AC ABP61917;
XX

DT 07-OCT-2002 (first entry)
DE Human lung cancer associated protein sequence SEQ ID NO:348.
KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
OS Homo sapiens.
PN WO200247534-A2.
PD 20-JUN-2002.
XX 30-NOV-2001; 2001WO-US047576.
PF 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX (CORI-) CORIXA CORP.
PA Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
PI WPI; 2002-583465/62.
DR N-P5DB; ABQ92440.
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX Example 2; Page 337-339; 381pp; English.
XX The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
CC invention
XX SQ Sequence 579 AA;
Alignment Scores:
Pred. No.: 6,17e-160 Length: 579
Score: 1875.00 Matches: 380
Percent Similarity: 84.68% Conservative: 51
Best Local Similarity: 74.66% Mismatches: 62
Query Match: 53.10% Indels: 16
DB: 5 Gaps: 9
US-09-270-437D-7 (1-1946) x ABP61917 (1-579)
QY 102 CCGGTTAGAGCGCGAAATTCAAATTCGAAATATTCACCCAGCTCCGATGGGAGTA 161
Db 77 ArgGlnArgIleArgLysLeuGlnIleArgAsnIleProProHisLeuTrpGluVal 96
QY 162 CTGGACAGCTGCTGCTCAGTATGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAG 221
Db 97 LeuAspSerLeuLeuValGlnIleGlyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGACGGCAGTGGTGAATGTCACTTATTCACACCGGGAGCAGACCGAGCAAGCATC 281
Db 117 SerGluThrAlaValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136

PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PA (CORI-) CORIXA CORP.
XX Wang T, Fan L;
XX
XX WPI; 2000-628399/60.
DR N-PSDB; AAC66035.
XX
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT protein is used for detecting and monitoring progression of lung cancer
PT in a patient.
XX
XX Claim 3; Page 259-261; 261pp; English.
XX
XX This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer
XX
XX Sequence 579 AA;
XX
XX Alignment Scores:
XX Pred. No.: 6,17e-160 Length: 579
XX Score: 1875.00 Matches: 380
XX Percent Similarity: 84.68% Conservative: 51
XX Best Local Similarity: 74.66% Mismatches: 62
XX Query Match: 53.10% Indels: 16
XX DB: 3 Gaps: 9
XX
XX US-09-270-437D-7 (1-1946) x AAB11365 (1-579)
XX
QY 102 CGGGTTCGAGCGCGGAAATTCGAATATTCACCCCGAGCTCCGATGGGAAGTA 161
Db 77 ArgGlnArgGlyLeuGlnLeuArgAenileProHisLeuGlnrPpGluVal 96
QY 162 CTGGACAGCCTGCTCGCTAGTATGTCAGTAGAGAACTGTGAGCAAGTGAAACCCGAG 221
Db 97 LeuAspSerLeuLeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAsp 116
QY 222 AGTGAGCGGAGCTGGTGAATGTCCTATTCACCGGGGAGCAGACGAGGCAAGCCATC 281
Db 117 SerGluThrAlaValAlaValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeu 136
QY 282 ATGAGCTGAATGGCCACGAGTGGAGAACCATGCCCTGAAGTCTCTTACATCCCGAT 341
Db 137 AspLysLeuAsnGlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrileProAsp 156
QY 342 GAGCAGATAGCA---CAGGGACCTGAGAATGGCGCGCCGAGGG-----GGCTTTGGCTCT 392
Db 157 GluThrAlaAlaGlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGln 176
QY 393 CGGGGTACGCCCGCCAGGCTCACCTGTGGCAGCGGGGGGCCCGCCAGCAGCAGCA 452
Db 177 ArgGlySerSerArgGlnGlySerPro-----GlySerValSerLysGlnLysPro 193
QY 453 GTGGACATCCCTTCGCTCGCTGTGGCCACCGCAGTATGTGGTCCCATTTATGGCAAG 512
Db 194 CysAspLeuProLeuArgLeuLeuValProThrGlnPheValGlyAlaIlelleGlyLys 213
QY 513 GAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGCTCCCAAGATAGACGTGCATAGG 572
Db 214 GluGlyAlaThrIleArgAsnilleThrLysGlnThrGlnSerLysIleAspValHisArg 233
QY 573 AAGGAGACGCGAGGTGCGACTGAAAGAGCCATCAGTGTGCATCCACCCCTGAGGGCTGC 632

Db 234 LysGluAsnAlaGlyAlaAlaGluLysSerIleThrIleLeuSerThrProGluGlyThr 253
QY 633 TCCTCCGCTTGAAGATGATCTTGAGATTATGATATAAGAGGCTTAAGACACCAACAG 692
Db 254 SerAlaAlaCysLysSerIleLeuGluIleMetHisLysGluAlaGlnAspIleLysPhe 273
QY 693 GCTGACGAGGTTCCCTCAGAGATCTGCGCCATATATCTTGTAGGGGCTCTCATTCG 752
Db 274 ThrGluGluIleProLeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGly 293
QY 753 AAGGAAGACGCAACCTGAAGAAGTAGACAGACATACCGAGACACAAAATCACCATCTCC 812
Db 294 LysGluGlyArgAsnLeuLysIleGluGlnAspThrAspThrLysIleThrIleSer 313
QY 813 TCGTTGCAAGACCTTACCTTTACACCTCGAGAGGACCATCTGTAAGGGGGCCATC 872
Db 314 ProLeuGlnGluLeuThrLeuTyrAsnProGluArgThrIleThrValLysGlyAsnVal 333
QY 873 GAGAAATTGTCAGGGCGGACGACGAAATATGAAGAAGTTTCGGAGGCTTATGAGAT 932
Db 334 GluThrCysAlaLysAlaGluGluIleMetLysLysIleArgGluSerTyrGluAsn 353
QY 933 GATGGGCTGCGCATGAGC-----TCTCACCTGATCCTCGGCTGAACCTGCTGCTGTA 986
Db 354 AspIleAlaSerMetAsnLeuGlnAlaHisIleIleProGlyLeuAsnLeuAlaLeu 373
QY 987 GGTCTTTTCCAGCTTCATCCAGCGCAGTCGCGCGC-----CCTCCAGCAGCGTT 1037
Db 374 GlyLeuPheProThrSerGlyMetProProThrSerGlyProProSerAlaMet 393
QY 1038 ACTGGGGTGTCTCCCTATAGCTCCCTTATGACGGCTCCCGACGAGATGTCGAGGTG 1097
Db 394 Thr-----ProProTyrProGlnPheGluGlnSer---GluThrGluThrValHisLeu 410
QY 1098 TTTATCCCGCCGAGGAGTGGCGCCATCATCGGCAAGAGGGGCGAGCAGATCAACAG 1157
Db 411 PheIleProAlaLeuSerValGlyAlaIlelleGlyLysGlnGlnHisIleLysGln 430
QY 1158 CTCTCCGCTTTCGAGCGCTCCATCAAGATTGACACCCGCAACACCTGACTCCAA 1217
Db 431 LeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLys 450
QY 1218 GTTCTATGTTTATCATCATCGACCGCGCAGGCGCCCAATTCAGGCTCAGGGAAGAATC 1277
Db 451 ValArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyAlaGln 470
QY 1278 TATGCAAACTCAAGGAGGAGAACTTCTTTGTCCTCCAGGAGAGTGAAGCTGGAGCC 1337
Db 471 TyrGlyLysIleLysGluGluAsnPheValSerProLysGluGluValLysLeuAla 490
QY 1338 CATATACGTGTCGACGATCAGCAGTCCCGGCTCATTTGGCAAAAGTGGAAAAACGGTG 1397
Db 491 HisIleArgValProSerPheAlaAlaGlyArgValIleGlyLysGlyLysThrVal 510
QY 1398 AACGATGTCGAAATTCAGCGCAGCTGAGTGTAGTAGTACCAAGAGACAGACCCCTGAT 1457
Db 511 AsnGluLeuGlnAsnLeuSerSerAlaGluAlaValProArgAspGlnThrProAsp 530
QY 1458 GAGAAACGACCGAGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCA 1517
Db 531 GluAsnAspGlnValValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGln 550
QY 1518 CGAGATCCGAGACATCTCTGGCCCGAGGTAAAGCAG---CAGCATCAGAGGGA---CAG 1571
Db 551 ArgLysIleGlnGluIleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGln 570
QY 1572 AGTAACCGCCCGCAGCAGGAGGAG 1598
Db 571 SerGlyProGlnSerArgArgLys 579
XX
XX RESULT 3
XX ABB75053

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:51:25 ; Search time 85.4096 Seconds
(without alignments)
12875.302 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 3531
Sequence: 1 gctgtagggagggtggg.....atttcctcaggttttaaaa 1946

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1596107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09270437/runat_23072004.095738.27403/app.query.fasta_1.5582
-DB=A Geneseq_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1 1 185 @runat_23072004.095738.27403 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database :
- 1: Geneseq_29Jan04.*
 - 2: Geneseqp1980s.*
 - 3: Geneseqp1990s.*
 - 4: Geneseqp2000s.*
 - 5: Geneseqp2001s.*
 - 6: Geneseqp2002s.*
 - 7: Geneseqp2003as.*
 - 8: Geneseqp2003bs.*
 - 9: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2491	70.5	577	2	AAY30649 A murine
2	1875	53.1	579	3	AAB11365 Human lun
3	1875	53.1	579	5	ABP75053 Human lun
4	1875	53.1	579	5	ABP74997 Human lun
5	1875	53.1	579	5	ABP75054 Human lun
6	1875	53.1	579	5	ABP61917 Human lun
7	1875	53.1	579	5	ABP61974 Human lun
8	1875	53.1	579	5	ABP61973 Human lun
9	1875	53.1	579	7	ADA28536 Recombina
10	1875	53.1	579	7	ADA28539 Recombina

11	1875	53.1	579	7	ADA28438 Human lun
12	1875	53.1	586	5	ABP75048 Human lun
13	1875	53.1	586	5	ABP61968 Human lun
14	1875	53.1	586	7	ADA28517 Recombina
15	1873	53.0	579	3	AAB11328 Human lun
16	1873	53.0	579	5	ABP74960 Human lun
17	1873	53.0	579	5	ABP61880 Human lun
18	1873	53.0	579	5	ABP61980 Human lun
19	1873	53.0	579	7	ADA28266 Human lun
20	1873	53.0	579	7	ADA14066 Human src
21	1802	51.0	619	4	ABG21963 Novel hum
22	1763.5	49.9	614	4	ABG06794 Novel hum
23	1763.5	49.9	620	4	AAB16163 Human nov
24	1763.5	49.9	620	6	ABU55232 Human nov
25	1740	49.3	583	4	ABG12592 Novel hum
26	1697	48.1	556	5	ABG96346 Human ova
27	1697	48.1	594	4	ABG06795 Novel hum
28	1585	44.9	319	4	AAM93826 Human pol
29	1573.5	44.6	555	6	ABU89799 Novel hum
30	1258	35.6	261	4	AAB16161 Human nov
31	1258	35.6	261	6	ABU55230 Human nov
32	1228	34.8	250	4	AAU16579 Human nov
33	1228	34.8	250	6	ABU55648 Human nov
34	863	24.4	558	4	ABBS8367 Drosophil
35	653	18.5	171	4	AAB16166 Human nov
36	605.5	17.1	187	6	AAO23971 Human IGF
37	510	14.4	192	6	AAU16164 Human nov
38	510	14.4	192	6	ABU55233 Human nov
39	502	14.2	171	4	AAU16583 Human nov
40	502	14.2	171	6	ABU55652 Human nov
41	465.5	13.2	148	4	ABG21962 Novel hum
42	410	11.6	266	4	ABG12593 Novel hum
43	409.5	11.6	209	4	ABG21961 Novel hum
44	292	8.3	93	4	AAM38501 Peptide #
45					

ALIGNMENTS

RESULT 1
AAY30649 standard; protein; 577 AA.
ID AAY30649 standard; protein; 577 AA.
XX
AC AAY30649;
XX
DT 17-NOV-1999 (first entry)
XX
DE A murine c-myc coding region determinant binding protein.
XX
DE c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
XX
KW endonucleolytic attack; half-life; breast cancer; colon cancer;
XX
KW pancreatic cancer.
XX
OS Mus musculus.
XX
XX
FN WO9946594-A2.
XX
PD 16-SEP-1999.
XX
PF 05-MAR-1999; 99WO-US004897.
XX
XX
PR 09-MAR-1998; 98US-0077372P.
XX
PR (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PA
XX
XX
PI Ross J;
XX
DR WPI; 1999-551506/46.
XX
DR N-PSDB; AAZ10617.
XX
PT Diagnosing presence or absence of a tumor in a human by examining c-myc
XX
PT coding region determinant-binding protein.

```
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4399556"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 88"
/notes="Organ: small intestine; Vector: pCMV-SPORT6;
Site 1: Not1; Site 2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH MGC Library."

ORIGIN

Query Match          34.0%; Score 662; DB 10; Length 896;
Best Local Similarity 93.4%; Pred. No. 1.8e-156;
Matches 727; Conservative 0; Mismatches 40; Indels 11; Gaps 3;

QY 821 AGACCTTACCTTTTACAACCTTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTG 880
Db 1 AGACCTTACCTTTTACAACCTTGAGAGGACCATCACTGTGAAGGGGCCATCGAGAATTG 60
QY 881 TTGACGGCCGACAGAGAAATATGAAGAAAGTTGCGGAGGCCTATGAGAAATGATGTGGC 940
Db 61 TTGACGGCCGACAGAGAAATATGAAGAAAGTTGCGGAGGCCTATGAGAAATGATGTGGC 120
QY 941 TGCATGAGC-----TCTCACCTGATCCCTGGCCTGAACCTGGCTGCTAGTCTCTTT 994
Db 121 TGCATGAGCCTGAGTCTCACTGATCCCTGGCCTGAACCTGGCTGCTAGTCTCTTT 180
QY 995 CCAGGTTTCAATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTA 1054
Db 181 CCAGGTTTCAATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTA 240
QY 1055 TAGCTCCTTTATGAGGCTCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCCCGCAGGC 1114
Db 241 TAGCTCCTTTATGAGGCTCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCCCGCAGGC 300
QY 1115 AGTGGCGCCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTGCAG 1174
Db 301 AGTGGCGCCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTGCAG 360
QY 1175 CGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGATGTTATCAT 1234
Db 361 CGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGATGTTATCAT 420
QY 1235 CACTGACCGCCAGAGCCCAATTCAAGCTCAGGAGAGATCTATGGCAACTCAAGGA 1294
Db 421 CACTGACCGCCAGAGCCCAATTCAAGCTCAGGAGAGATCTATGGCAACTCAAGGA 480
QY 1295 GGAGAACTTTTGTGTCCTCAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGC 1354
Db 481 GGAGAACTTTTGTGTCCTCAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGC 540
QY 1355 ATCAGCAGCTGGCCGGGTGATTGGCAAGGTGAAAAACGGTGAAACGAGTTGCAGAAATT 1414
Db 541 ATCAGCAGCTGGCCGGGTGATTGGCAAGGTGAAAAACGGTGAAACGAGTTGCAGAAATT 600
QY 1415 GACGGCAGCTGAGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGGTGAT 1474
Db 601 GACGGCAGCTGAGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACAGGTGAT 660
QY 1475 CGTGAAG-----ATCATCGGACATTTCTATGCCAG-TCAGATGGCTCAACGGAGATCCGA 1529
Db 661 CGTGAAGACATCCATCCGGAACATTTCTATGCCAGCTCAGATGGCTCAGGAATCCGAG 720
QY 1530 GACATCCTGGCCCGAGTTAAGCAGCAGCATCAAGAGGGAAGAGTAACAGGCCCGCAGG 1587
Db 721 AACTTCTGGCCCGAGTTAAGCAGCAGCATCAAGAGGACGGGTTCCCGCGCGCGG 778
```

Search completed: July 24, 2004, 16:42:21
Job time : 4881.33 secs

Db	481	TTCCAGTGAAGAACCTTGATCTCTCAGCCCAACACACCCAAATTTGGCCCACTGT	540
Qy	1809	NTGCCCTCCGGGTGTGAGAAATTTAGCGAAGGCACTTTTAAACGTGGATTTTAAA	1868
Db	541	CTGCCCCCTCGGGTGTGAGAAATTTCTAGCGAAGGCACTTTTAAACGTGGATTTAAA	600
Qy	1869	GAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCAACCTCAGTGGGAAGAAAATAAAT	1928
Db	601	GAAGCTCTCCAGGCCCCACCAAGAGGGTGGATCAACCTCAGTGGGAAGAAAATAAAT	660
Qy	1929	TTCCCTCAGGTTTT	1942
Db	661	TTCCCTCAGGTTTT	674
RESULT 14			
LOCUS	B0651665	955 bp	mRNA
DEFINITION	AGENCOURT 8297761 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6269842		EST 15-JUL-2002
ACCESSION	B0651665		
VERSION	B0651665.1	GI:21775837	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: CGAP (Stanford) cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LUCM2444 row: a column: 11 High quality sequence stop: 588. Location/Qualifiers		
FEATURES			
source	1..955	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:6269842"	
		/tissue type="hepatocellular carcinoma, cell line"	
		/lab host="DH10B (phage-resistant)"	
		/clone lib="NIH MGC 100"	
		/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN			
Query Match	34.0%;	Score 662.2;	DB 13; Length 955;
Best Local Similarity	98.8%;	Pred. No. 1.7e-156;	
Matches 667;	Conservative 0;	Mismatches 8;	Indels 0; Gaps 0;
Qy	264	CAGACGAGGCAAGCCATCATGAAGCTGAATGCCACCACTTGGAGAACCATGCCCTGAAG	323
Db	1	CAGACGAGGCAAGCCATCATGAAGCTGAATGCCACCACTTGGAGAACCATGCCCTGAAG	60
Qy	324	GTCTCTACATCCCGATGACGATAGACACAGGACCTGAGATGGCGCGCGAGGGGC	383

Db	61	GTCTCTACATCCCGATGACGATAGACACAGGACCTGAGAATGGCGCGCGAGGGGC	120
Qy	384	TTTGGCTCTCGGGTGAAGCCCGCCGCGGCTCACTGTGCGAGCGGGGCCCGCCGCAAG	443
Db	121	TTTGGCTCTCGGGTGAAGCCCGCCGCGGCTCACTGTGCGAGCGGGGCCCGCCGCAAG	180
Qy	444	CAGCAGCAAGTGGACATCCCTTCGGCTCTCTGTGTGCCACCCAGTATGTGGGTGCCATT	503
Db	181	CAGCAGCAAGTGGACATCCCTTCGGCTCTCTGTGTGCCACCCAGTATGTGGGTGCCATT	240
Qy	504	ATTGGCAAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGAC	563
Db	241	ATTGGCAAGAGGGGGCCACCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGAC	300
Qy	564	GTGCATAGGAAGAGAAACGACAGGTGCGAGTGAAGAGCCATCAGTGTGCATCCACCCCT	623
Db	301	GTGCATAGGAAGAGAAACGACAGGTGCGAGTGAAGAGCCATCAGTGTGCATCCACCCCT	360
Qy	624	GAGGGCTGCTCCTCGGCTTGTAAAGATGATCTTGAGATTATGTCATAAGAGGCTAAGGAC	683
Db	361	GAGGGCTGCTCCTCGGCTTGTAAAGATGATCTTGAGATTATGTCATAAGAGGCTAAGGAC	420
Qy	684	ACCAAAACGGCTGACGAGGTTCCCTCTGAAGATCTCTGGCCCATATAACTTTGTAGGGCGT	743
Db	421	ACCAAAACGGCTGACGAGGTTCCCTCTGAAGATCTCTGGCCCATATAACTTTGTAGGGCGT	480
Qy	744	CTCATTTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAATC	803
Db	481	CTCATTTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAATC	540
Qy	804	ACCATCTCCTCGTTGCAAGACCTTACCCCTTACAAACCTTGAGAGACCATCACTGTGAAG	863
Db	541	ACCATCTCCTCGTTGCAAGACCTTACCCCTTACAAACCTTGAGAGACCATCACTGTGAAG	600
Qy	864	GGGGCCATCGAGAATTGTTGAGGGCCGAGCAGGAAATTAATGAAGAAAGTTGGGAGGCC	923
Db	601	GGGGCCATCGAGAATTGTTGAGGGCCGAGCAGGAAATTAATGAAGAAAGTTGGGAGGCC	660
Qy	924	TATGAGAAATGATGTG	938
Db	661	TATGCAATGACGTG	675
RESULT 15			
LOCUS	BF984962	896 bp	mRNA
DEFINITION	602308484F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4399556 5', mRNA sequence.		linear EST 23-JAN-2001
ACCESSION	BF984962		
VERSION	BF984962.1	GI:12387774	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 896)		
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LILAM10103 row: h column: 21 High quality sequence stop: 652. Location/Qualifiers		
FEATURES			
source	1..896	/organism="Homo sapiens"	


```
QY 1051 CCTATAGCTCTTTATGAGGCTCCGAGCAGGAGATGCTGACGCTGTTTATCCCGCCC 1110
Db 1448 CTCCTTACCACAAATTTGAGCAATCAGAGACGGAGACTGTGCACTCTTTATCCCGCCC 1507
QY 1111 AGCAGTGGGGCCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGTTTG 1170
Db 1508 TGTCCGTTGGCCCATCATTTGGCAAGCAGGCGCAACATCAACACAGCTTCTCGCTTG 1567
QY 1171 CCAGCGCTTCATCAAGATTGACACACCGGAAACACCTGACTCCAAAAGTTCGTATGTTA 1230
Db 1568 CGGAGCTTCGATTAAAGATCGCTCCAGCAGAAGCACCAGATGCTAAAAGTGGGATGTA 1627
QY 1231 TCATCACTGAGCGCCAGAGGCCCAATTCAGGCTCAGGAGAGATCTATGCGAACTCA 1290
Db 1628 TTATCACTGAGCACCAGAGGCTCAGTTCAAGGCTCAGGAGAGAAATTTATGGAATAA 1687
QY 1291 AGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATAGTGTGC 1350
Db 1688 AAGAAGAAACTTTGTTAGTCTTAAAGAGAGGTGAACTTGAAGCTCAATCAGATGC 1747
QY 1351 CAGCATCAGAGCTGGCGGGTCAATTGGCAAGGTGAAAGACGGTGAACGAGTTGCAGA 1410
Db 1748 CGTCTTGTCTGCTGGCAGAGTTATTGGGAAAGAGGACAAACCGTGAATGAGCTCCAGA 1807
QY 1411 ATTGACGGCAGCTGAGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAAACGACAGG 1470
Db 1808 GTTATCAAGTGTGAAGTTGTGTCTCCCGTGACCCAGACACCTGATGAGAAATGATCAAG 1867
QY 1471 TCATCGTGAATAATCATCGGACATTTCTATCCAGTCAGATGGCTCAACGGAAGATCCGAG 1530
Db 1868 TAGTTGCAAAATAAATCGGCCACTTCTATCTGTTGCCAGGTTGCCAGAGGAATTCAGG 1927
QY 1531 ACATCTGGCCGAGGTTAAGCAGCAGCATCAGAAG 1565
Db 1928 AAATCTGACTAGGTAAGCAGCAGCAGCAGCAG 1962
```

RESULT 10

B0647360

LOCUS B0647360 953 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT_8443628 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6284231
5', mRNA sequence.

ACCESSION B0647360

VERSION B0647360.1 GI:21771532

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgs.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: CGAP (Stanford)

cDNA Library Preparation: Ruben Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnl.gov>

Plate: LLC2481 row: h column: 24

High quality sequence stop: 651.

Location/Qualifiers

1..953

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6284231"

/tissue_type="hepatocellular carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

FEATURES

source

/clone_lib="NIH_MGC_100"
/notes="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

```
Query Match 34.9%; Score 678.6; DB 13; Length 953;  
Best Local Similarity 95.1%; Pred. No. 1.1e-160;  
Matches 765; Conservative 0; Mismatches 30; Indels 9; Gaps 6;  
QY 108 AGGAGCGGGAATTCAAATCCGAAATATTCCACCCCGAGCTCCGATGGGAGTACTGGAC 167  
Db 138 AGGAGCGGGAATTCAAATCCGAAATATTCCACCCCGAGCTCCGATGGGAGTACTGGAC 197  
QY 168 AGCTGCTGGCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227  
Db 198 AGCTGCTGGCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 257  
QY 228 ACGGCAAGTGGTGAATGTCACCTATTTCACACCGGAGCAGACCCAGGCAAGCCATCATGAAG 287  
Db 258 ACGGCAAGTGGTGAATGTCACCTATTTCACACCGGAGCAGACCCAGGCAAGCCATCATGAAG 317  
QY 288 CTGAATGGCCACCCAGTTGGAGAACCATGCCCTCAAGGTCTCTCTACATCCCCGATGAGCAG 347  
Db 318 CTGAATGGCCACCCAGTTGGAGAACCATGCCCTCAAGGTCTCTCTACATCCCCGATGAGCAG 377  
QY 348 ATAGCACAGGACCTCAGAAATGGCGCCCGAGGGGGCTTTTGGCTCTCGGGGTCAAGCCCGCG 407  
Db 378 ATAGCACAGGACCTCAGAAATGGCGCCCGAGGGGGCTTTTGGCTCTCGGGGTCAAGCCCGCG 437  
QY 408 CAGGGCTCACCTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 467  
Db 438 CAGGGCTCACCTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 497  
QY 468 CGGCTCCTGGTGGCCACCCAGTATGTTGGTGCCTATTATTTGGCAAGAGGGGGGCCACCATC 527  
Db 498 CGGCTCCTGGTGGCCACCCAGTATGTTGGTGCCTATTATTTGGCAAGAGGGGGGCCACCATC 557  
QY 528 CGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGATAGGAGGAGAGAACGAGGT 587  
Db 558 CGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGATAGGAGGAGAGAACGAGGT 617  
QY 588 GCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCGCTTGTAAAG 647  
Db 618 GCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCGCTTGTAAAG 677  
QY 648 ATGATCTTGGAGATTATGATAAAGAGGTAAGGACACCAAAA-CGGCTGACAGAGTTCC 706  
Db 678 ATGATCTTGGAGATTATGATAAAGAGGTAAGGACACCAAAAACCGCTGACGAGTTCC 737  
QY 707 CTGTAAGATCTTGGCCCATATAAAC-TTTGTAGGGGTCTCATTTGGCAAGGAGAGACGGA 765  
Db 738 CTGTAAGATCTTGGCCCATATAAACCTTTTGTAGGGGTCTCATTTGGCAAGGAGAGACGGA 797  
QY 766 A-CCTGAGAGAGGTAGAGCAAGATA-CCGAGACAAAATCACCATCTCTCGTT-GCAAG 822  
Db 798 ACCCTGAAGAGGTANAGCCAGATACCCGAGACAAAATCACCATCTCTCTCGTTGGCAGA 857  
QY 823 ACCTTACCTTTTACAAACCTG----AGAGGACCATCTACTGTGAAGGGGCCCATCGAGAT 878  
Db 858 ACCTTACCTTTTACAAACCTTGGAGAGGACATCCACTGTGAAGGGGCCCATCGAGAT 917  
QY 879 TGTTCAGGCGCGAGCAGGAATA 902  
Db 918 TGTTCAGGCGCGCGAGGAAA 941
```



```

Db      1846 TAGTGTCAAATAAATGCGCACTTCTATGTTGCCAGGTTGCCAGAGGAATTCAGG 1905
QY      1531 ACATCTGGCCCGAGGTTAAGCAGCAGATCAGAAG 1565
Db      1906 AAATTCTGACTCAGGTAAGCAGCAGCAGCAGCAG 1940

RESULT 8
LOCUS   BX327672/c
DEFINITION BX327672 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DC018Y118 5-PRIME, mRNA sequence.
ACCESSION BX327672
VERSION   BX327672.1 GI:30342696
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 919)
          Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
          Full-length cDNA libraries and normalization
          Unpublished (2001)
COMMENT   Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          Library was constructed by life Technologies, a division of
          Invitrogen. This sequence belongs to sequence cluster 8684.r For
          more information about this cluster, see
          http://www.genoscope.cns.fr/
          cgi-bin/cluster.cgi?seq=CS0BAG007ZH09_CS00648_1&cluster=8684.r.
          Contact : Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CS0BAG007ZH09_CS00648_1.
FEATURES
         Location/Qualifiers
             1..919
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="CS0DC018Y118"
                 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
                 primer. Five prime end enriched, double-strand cDNA was
                 digested with Not I and cloned into the Not I and EcoR V
                 sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 37.1%; Score 722.6; DB 13; Length 919;
Best Local Similarity 98.4%; Pred. No. 7.8e-172;
Matches 750; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY      108 AGAGCGGGAATTCGAATTCG-AAATATTCACCCCGAGCTCCGATGGGAAGTACTGGA 166
Db      783 AGGACCGGGAATTCGAATTCGAAATATTCACCCCGAGCTCCGATGGGAAGTACTGGA 724
QY      167 CAGCTGTGGCTCAGTATGGTACAGTAGAAGTCTGAGCAAGTGAACACCGAGTGGA 226
Db      723 CAGCTGTGGCTCAGTATGGTACAGTAGAAGTCTGAGCAAGTGAACACCGAGTGGA 664
QY      227 GACGGCAGTGGTGAATGTCACTTATTCACCGGGAGCAGACCCAGCAAGCCATCATGAA 286
Db      663 GACGGCAGTGGTGAATGTCACTTATTCACCGGGAGCAGACCCAGCAAGCCATCATGAA 604
QY      287 GCTGAATGCCACCAAGTGGAGAACCATGCCCTGGAAGGCTCTCTACATCCCGATGAGA 346
Db      603 GCTGAATGCCACCAAGTGGAGAACCATGCCCTGGAAGGCTCTCTACATCCCGATGAGA 544
QY      347 GATAGCACAGGACCTGAGAAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTACGCCCG 406
Db      543 GATAGCACAGGACCTGAGAAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTACGCCCG 484

```

```

QY      407 CAGGGCTACCTGTGGCAGCGGGGGCCCCAGCCAAAGCAGCAGAAAGTGGACATTCGCCCT 466
Db      483 CAGGGCTACCTGTGGCAGCGGGGGCCCCAGCCAAAGCAGCAGAAAGTGGACATTCGCCCT 424
QY      467 TCGGCTCTCTGGTCCCAACCCAGTATCTGGTGCATTATTATTTGGCAAGGAGGGGGCCACCAT 526
Db      423 TCGGCTCTCTGGTCCCAACCCAGTATCTGGTGCATTATTATTTGGCAAGGAGGGGGCCACCAT 364
QY      527 CCGCAACATCACAAAAACAGACCCAGTCCAAAGATAGAGCTGCATAGGAAGGAGAACGCGAG 586
Db      363 CCGCAACATCACAAAAACAGACCCAGTCCAAAGATAGAGCTGCATAGGAAGGAGAACGCGAG 304
QY      587 TGCAGCTGAAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTCTCCGCTTGTA 646
Db      303 TGCAGCTGAAAAAGCCATCAGTGTGCATCTCCACCCCTGAGGGCTGCTCTCCGCTTGTA 244
QY      647 GATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGGTTC 706
Db      243 GATGATCTTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGGTTC 184
QY      707 CTGGAAGATCTCGGCCCAATAAATCTTTAGGGCTCTCTATTGGCAAGGAGACGCGAA 766
Db      183 CTGGAAGATCTCGGCCCAATAAATCTTTAGGGCTCTCTATTGGCAAGGAGACGCGAA 124
QY      767 CTGGAAGAGGTAGACAGATACCCAGACCAAAATCACAATCTCTCTGTTGCAAGACCT 826
Db      123 CTGGAAGAGGTAGACAGATACCCAGACCAAAATCACAATCTCTCTGTTGCAAGACCT 64
QY      827 TACCCCTTTACAAACCTTGAGAGGACCATCACTGTGAAGGGGGC 868
Db      63 TACCCCTTTACAA-CTTGAGAGGACCATCACTGTGAAGGGGGC 23

RESULT 9
LOCUS   AK011689
DEFINITION Mus musculus 2202 bp mRNA linear HTC 20-SRP-2003 enriched library, clone:2610036H18 product:insulin-like growth factor 2, binding protein 3, full insert sequence.
ACCESSION AK011689.1 GI:12847972
VERSION   AK011689
KEYWORDS HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
          Carninci, P. and Hayashizaki, Y.
          High-efficiency full-length cDNA cloning
          Meth. Enzymol. 303, 19-44 (1999)
          99279253
          10349636
REFERENCE 2
          Carninci, P. and Hayashizaki, Y.
          Normalization and subtraction of cap-trapper-selected cDNAs to
          Prepare full-length cDNA libraries for rapid discovery of new genes
          Genome Res. 10 (10), 1617-1630 (2000)
          20499374
          11042159
REFERENCE 3
          Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
          Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
          Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
          Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
          Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
          Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
          Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
          RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multipillar sequencer
          Genome Res. 10 (11), 1757-1771 (2000)
          20530913
          11076861

```

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

source
1..4100

/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="PANTOM_DB:E430018F06"
/db_xref="MGI:247362"
/db_xref="taxon:10090"
/clone="E430018F06"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
243..1982

CDS
/note="unnamed protein product; insulin-like growth factor
2, binding protein 3 (MGI:MGI:1890359, GB|NM_023670,
evidence: BLASTN, 99%, match=1740)
putative"

/codon_start=1
/protein_id="BAC40370.1"
/db_xref="GI:26353480"
/translation="MNKLIYIGNLSDHAGPADLESVFKDAKIPVAGPFLVITGYAFVDC
PDEGWLKALSGKMLHGKPMVEHSPKQRIKQIRNPPLHQLWVLDLIV
QYGVSECVNTDSEAVNVTVSSKQARQALDKLNGFOLENTLKVAIPIPETAA
QONSPQLRGRGQRGSRQASPGSVSKPCDPLRLIVPTQFVCAIIGKEGATI
RNIITQGTQKIDVHKENTGAERKSIITILSPGTSARCKSIILMKKEAQDIKTEE
IPLKILAHNNVPLGKGNLKKIEQDITKITSPLQELTLNPERTLVKGSVE
TCARKEEIMKIRESENDIASMNLQANLIPGNLNLGFLPPTSGMPPTSPFPT
LTPYPOFEQSETVHLFIPALSGAIIGQGQHIKOLSRFAGASIKIAPABAPDAK
VRMIIITGPPPAQKAQRIYVKIKEENFVSPKEVLEAHIRVPSFAAGRVIKGGK
TVNLSLSLSVAVVPRDQTPDENDQVVVKITGHFYACQVAQRKIQELITQVKHQHQ
KALQSGFPQSRKK"
4079..4084
/note="putative"
4100
/note="putative"

polyA_signal

4079..4084

polyA_site

4100

ORIGIN

Query Match 37.2%; Score 723.4; DB 11; Length 4100;
Best Local Similarity 68.7%; Pred. No. 1.1e-171;
Matches 1013; Conservative 0; Mismatches 456; Indels 6; Gaps 1;
QY 97 CTTCCCGGTTAGGAGCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGG 156
DB |||||
QY 466 CTAAACGGACAGGATTCTGTAACCTTCAGATACGAAATATCCGCCCACTTACAAATGGG 525
DB |||||
QY 157 AAGTACTGGACAGCGCTGCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACA 216
DB |||||
QY 526 AGGTCTGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACA 585
DB |||||
QY 217 CCGAGAGTGGACGCGAGTGGTGAATGTCACTATTCCAACCGGGAGCGAGCAGGCAAG 276
DB |||||
QY 586 CGGATTCGAAACGGCAGTTGTAATGTAACTATTCAGTAGAAGGACCAAGCTAGACAG 645
DB |||||
QY 277 CCATCATGAGCTGAATGGCCACAGTGTGGAAACCATGCCCTGAAGGTCTCTTACATCC 336
DB |||||
QY 646 CTTAGACAAACTGAATGATTCACAGTTAGAGAACTTACCTTGAAGTTGCTTACATCC 705
DB |||||
QY 337 CCGATGAGCAGATAGCAGGAGCACTGGAATGGGCGCGAGGGGCTTTGGCTCTCGGG 396
DB |||||
QY 706 CAGATGAACCTCTGCCAGCAAAATCCCTCCACACAGCTCCGGGGCGCGGGGCCAG 765
DB |||||
QY 397 GTACAGCCCGCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGAGTGG 456

Db 766 GGCAGCGGGGCTCATCCAGGCGAGCGCTCTCCAGGATCGGTGTCCAAGCAGAAACCTCTGT 825
QY |||||
QY 457 ACATCCCCCTTTCGGCTCCTGTGTGCCACCCAGTAGTATGGGTGCCATTTATTCGCAAGGAGG 516
Db |||||
Db 826 ACCTGCCACTGGCGCTGTGGTTCCTCCACCCAGCTTTGTGTGGAGCCATTATAGGAAAGAG 885
QY |||||
QY 517 GGGCCACATCCGCAACATCACAACACAGACCCAGTCCAAAGATAGACCTGTGATAGGAGG 576
Db |||||
Db 886 GTGTACTACATTCGCAACATCACAAGCAGACCCAGTCTAAATTCATGATCGATCGATGAAG 945
QY |||||
QY 577 AGAACGCGAGGTGACGTGAAAAGCCATCAGTGTGCACTCCACCCCTCAGGCGCTCTCTCT 636
Db |||||
Db 946 AGAATACAGGGCGCGGAGAGTCCATTACTCTCTCTACCCCTCAGGSCACCTCTG 1005
QY |||||
QY 637 CCGCTTTGTAAGATCATCTTGAGATATATGCAATAAGAGGCTTAAGGACACCAAAACGGCTG 696
Db |||||
Db 1006 CAGCTTGTAACTCTATTCTGGAGATTATGCAATAAGGAAGCTCAAGATATAAATTCACAG 1065
QY |||||
QY 697 ACGAGTTCCCTCGAGATCCTGCCCATATACTTTGTAGGGCTCTCATTTGGCAAGG 756
Db |||||
Db 1066 AGGAGATTCCCTTGAAGATCTTAGCTCAATAAATCTTTGTAGGCGCTCTCATTTGTAAG 1125
QY |||||
QY 757 AAGCACGGAACCTGAAAGAGGTAGAGCAAGATACCGAGACAAAATACCATCTCTCTCT 816
Db |||||
Db 1126 AAGGAGAAACCTTAAATAAATCGAGCAAGACACGACACTAAATACAAATATCTCCAT 1185
QY |||||
QY 817 TGCAGACCTTACCTTTACAACTTGAGAGACCATCATCTGTGAAGGGGCCATCGAGA 876
Db |||||
Db 1186 TGCAGGAACCTGACGCTGTACCAATCCGGAACGACCACTTACAGTGAAGAGCAGTGTGAGA 1245
QY |||||
QY 877 ATTGTTGAGGGCGGAGCAGCAAGAAATATGAAGAAGTTCGGGAGGSCCTATCAGAAATGATG 936
Db |||||
Db 1246 CGTGTGCCAAGCGGAGAGGAAATATGAAGAAGATCAGGAGGCTTATGAAAATGATA 1305
QY |||||
QY 937 TGGCTGCCATCAGCTCTC-----ACCTGATCCCTGGCGCTGAAACCTGGCTGTGTAGGTC 990
Db |||||
Db 1306 TTGCTTCCTGAACTTCAAGCAAAATTTAACTCCCTGGATTAATCTGATGCTTGGGTC 1365
QY |||||
QY 991 TTTTCCAGCTTCATCCAGCGAGCTCCGCGCGCTCCCGAGAGCTTATCTGGGGCTGCTC 1050
Db |||||
Db 1366 TGTTCACCCAGCTCAGGATGCCACTCCACCTCAGGCGCCCTTCAACCCCTGACTC 1425
QY |||||
QY 1051 CTTATAGCTCTTTATGAGGCTCCGAGCAGGAGATGGTGCAGTGTATTATCCCGGCC 1110
Db |||||
Db 1426 CTCCTACCCCAATTTGAGCAATCAGAGACGAGACTGTGTGATCTGTATTATCCGCC 1485
QY |||||
QY 1111 AGGCAGTGGGCGCCATCATCGGCAAGAGGGGCGACACATCAAAAGCTCTCCCGTTTG 1170
Db |||||
Db 1486 TGTCCGTTGGCGCCATCATTCGCAAGCAGGGCCACACATCAACAGCTTCTCGCTTTG 1545
QY |||||
QY 1171 CCAGCGCTTCCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTCGTATGGTTA 1230
Db |||||
Db 1546 CGGAGCTTCGATTAAAGATCGCTCCAGCGCAAGCACCAGATCTAAAGTCGGGATGGTGA 1605
QY |||||
QY 1231 TCATCAGTGGACCGCGAGGCGCCAAATTCAGGCTCAGGGAAGATCTATGCAAACTCA 1290
Db |||||
Db 1606 TTATCAGTGGACCAAGAGGCTCAGTTCAGGCTCAGGGAAGAAATTTATGAAAATTA 1665
QY |||||
QY 1291 AGGAGGAGAACTTCTTTGGTCCAAAGGAGGAAGTGAAGCTGGAGAGCCACATACGTTGC 1350
Db |||||
Db 1666 AAGAGAAACCTTTGTAGTCTTAAAGAGAGTGAACCTTGAAGCTCACATCAGAGTGC 1725
QY |||||
QY 1351 CAGCATCAGAGCTGGCGGGTTCATTTGGCAAGGTGGAAAACCGTGAACGAGTTCAGAG 1410
Db |||||
Db 1726 CGTCTCTTCTCTGGCAGAGTTATTGGGAAAGAGGAGGCAAAACCGTGAATAGCTCCAGA 1785
QY |||||
QY 1411 ATTTGAGCGCAGCTCAGTGTGTAGTACCAAGAGACAGACCCCTCTGATGAGAGACGACAGG 1470
Db |||||
Db 1786 GTTTATCAAGTCTGAAGTTGTGTCCTCCCGGTGACCAAGACACTGATGAGATGATCAAG 1845
QY |||||
QY 1471 TCATCGTGAATAATCATCGACACTTTCTATGTCAGTCAAGTGGCTCAACGGAAGATCCGAG 1530
Db |||||

AK088465	LOCUS	4100 bp	linear	HTC 20-SEP-2003
Mus musculus 2 days neonate thymic cells cDNA, RIKEN full-length enriched library, clone:E430018P06 product:insulin-like growth factor 2, binding protein 3, full insert sequence.				
AK088465	ACCESSION	GI:26353479		
AK088465.1	VERSION			
HTC; CAP trapper.	KEYWORDS			
Mus musculus (house mouse)	SOURCE			
Mus musculus	ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
1	REFERENCE			
Carninci, P. and Hayashizaki, Y.	AUTHORS			
High-efficiency full-length cDNA cloning	TITLE			
Meth. Enzymol. 303, 19-44 (1999)	JOURNAL			
99279253	MEDLINE			
10349636	PUBMED			
2	REFERENCE			
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	AUTHORS			
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	TITLE			
Genome Res. 10 (10), 1617-1630 (2000)	JOURNAL			
20499374	MEDLINE			
11042159	PUBMED			
3	REFERENCE			
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	AUTHORS			
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	TITLE			
Genome Res. 10 (11), 1757-1771 (2000)	JOURNAL			
20530913	MEDLINE			
11076861	PUBMED			
4	REFERENCE			
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium	AUTHORS			
Functional annotation of a full-length mouse cDNA collection	TITLE			
Nature 409, 685-690 (2001)	JOURNAL			
5	REFERENCE			
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	AUTHORS			
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	TITLE			
Nature 420, 563-573 (2002)	JOURNAL			
6 (bases 1 to 4100)	MEDLINE			
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanganaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	AUTHORS			
Direct Submission	TITLE			
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	JOURNAL			
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken				

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1042)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12263 row: 1 column: 24
High quality sequence stop: 682.
Location/Qualifiers
1. 1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5550887"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

FEATURES
source
1. 1042

ORIGIN
Query Match 42.6%; Score 828.4; DB 12; Length 1042;
Best Local Similarity 96.8%; Pred. No. 1.2e-198;
Matches 920; Conservative 0; Mismatches 18; Indels 12; Gaps 7;
676 CTAGGACACCAAAACGGCTGACGAGTTCCCTGAAGATCTCGCCCATTAATCACTTTG 735
2 CTCAGGACACCAAAACGGCTGACGAGTTCCCTGAAGATCTCGCCCATTAATCACTTTG 61
736 TAGGGCTCTCATTTGGCAGGAAGGACGGAACCTGAGAGGTAGAGCAATACCGAGA 795
62 TAGGGCTCTCATTTGGCAGGAAGGACGGAACCTGAGAGGTAGAGCAATACCGAGA 121
796 CAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACAACTCAGAGGACCATCA 855
122 CAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACAACTCAGAGGACCATCA 181
856 CTGTGAAGGGGGCCATCAGAAATTTGTTCAGGGCCGAGCAGGAATATGAAGAAATTC 915
182 CTGTGAAGGGGGCCATCAGAAATTTGTTCAGGGCCGAGCAGGAATATGAAGAAATTC 241
916 GGGAGCCCTATGAGAAATGATGTGGTGCATGAGC-----TCTCAGCTGATCCCTGGCC 969
242 GGGAGCCCTATGAGAAATGATGTGGTGCATGAGC-----TCTCAGCTGATCCCTGGCC 301
970 TGAACCTGGCTGTAGTCTTTTCCAGCTTCAATCAGGCGAGTCCCGCGCCCTCCCA 1029
302 TGAACCTGGCTGTAGTCTTTTCCAGCTTCAATCAGGCGAGTCCCGCGCCCTCCCA 361
1030 GCAGGTTACTGGGGCTCTCCCTATAGTCTCTTTATGACGCTCCCGAGCAGGAGATGG 1089
362 GCAGGTTACTGGGGCTCTCCCTATAGTCTCTTTATGACGCTCCCGAGCAGGAGATGG 421
1090 TGCAGGTTACTGGGGCTCTCCCTATAGTCTCTTTATGACGCTCCCGAGCAGGAGATGG 1149
422 TGCAGGTTACTGGGGCTCTCCCTATAGTCTCTTTATGACGCTCCCGAGCAGGAGATGG 481
1150 TCAACAGCTCTCCCGGTTTGCAGGCTTCCATCAAGATTCACACCCGAGGAGGAGGAG 1209
482 TCAACAGCTCTCCCGGTTTGCAGGCTTCCATCAAGATTCACACCCGAGGAGGAGGAG 541

QY 1210 ACTCCAAAGTTCTGATGTTATCATCTACTGACCGCCAGAGGCCCAATTTCAAGGCTCAGG 1269
Db 542 ACTCCAAAGTTCTGATGTTATCATCTACTGACCGCCAGAGGCCCAATTTCAAGGCTCAGG 601
QY 1270 GAGAAATCTATGGCAAACTCAAGAGGAGAACTTTCTTTTGGTCCCAAGAGGAGTGAAGC 1329
Db 602 GAGAAATCTATGGCAAACTCAAGAGGAGAACTTTCTTTTGGTCCCAAGAGGAGTGAAGC 661
QY 1330 TGGAGACCCACATACGTTGTCAGCATCAGCAGTGGCCGGTCAATTTGGCAAGGTTGAA 1389
Db 662 TGGAGACCCACATACGTTGTCAGCATCAGCAGTGGCCGGTCAATTTGGCAAGGTTGAA 721
QY 1390 AAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTGTAGTACCAAGAGACCA 1449
Db 722 AAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTGTAGTACCAAGAGACCA 781
QY 1450 CCCCTGATGAGAACGACCATGTCATCGTGAATCATC-GGACATTTCTATGCCAGTCAG 1508
Db 782 CCCCTGATGAGAACGACCATGTCATCGTGAATCATC-GGACATTTCTATGCCAGTCAG 841
QY 1509 ATGCTCAAC-GGAAGATCCAGACATCTT-GGCCAGGTTAAGCAGCAGCAT-CAGAG 1565
Db 842 ATGCTCAACGGAAGATCCAGACATCTTGGGGCCAGGTTAAGCAGCAGCATCAAAA 901
QY 1566 GGACAGA-GTAACCCAGGCCAGGCAC-GGAGGAAGTACACGAGCCCTCC 1613
Db 902 GGACAGAGGTAACCCAGGCCAGGCACGGAGGAAGTACACGAGCCCTCC 951

RESULT 6
LOCUS BQ648383 950 bp mRNA linear EST 15-JUL-2002
DEFINITION AGENCOURT 8342265 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6268202
5', mRNA sequence.
ACCESSION BQ648383
VERSION BQ648383.1 GI:21772555
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 950)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2439 row: m column: 03
High quality sequence stop: 599.
Location/Qualifiers
1. 950
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6268202"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 100"
/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC

FEATURES
source
1. 950

Db

1893

ATCATCGGACATTTCTATGCGAGCCAGATGGCTCAGCGGAAGATCCGAGACATCTCGCT

1952

Qy

1542

CAGGTTAAGCAGCAGCATCAAGAGGACAGAGTAACCAAGGCCAGAGGCCAGGCGGACGAGGAAGTGA

1601

Db

1953

CAAGTTTAAAGCAAGCAGCAGGAGGAGAGCAACCTGGCCAGGCTCGAGGAAGTGA

2012

Qy

1602

-CCAGCCCTCTCTCTCCCTTNGAGTCCAGGACAAACCGGCGAGAA-----

1647

Db

2013

CCCCGCCCTCTCTGCTCCATGCTCCAGATCAGCAGGAGGAACACAGAACTGGAGG

2072

Qy

1648

-----ATCAGAGTGTGCTCTCCCGGAGGCTCAGATGAGTGGGAATCCGGAC

1699

Db

2073

CGCGGTGGAGGCCGCTGTCTCTTCCAGCAGGCTCAGATGAGTGGGAATCAGG-C

2131

Qy

1700

ACNTGGCGCGGCTGTAGATCAGGTTTCCCACTTGATTGAGAAAGATGTTCCAGTGAGG

1759

Db

2132

ATTGGGCTGCTGGAGATCAGGTTTGCACTGCTTGAGAACAAATGTTCCAGTGAGG

2191

Qy

1760

AACCTGTATCTNTCAGCCCCAAACACCCCAATGTCGCCAAACACTGNTGCCCTCGG

1819

Db

2192

AATCTGTATCTCTCGCCCCCAA--TTGAGCCAGCTGGCCACAGCCCCACCTTGGAAATAT

2249

Qy

1820

GGTGTCAAAATTTAGCGCAAGCAGCACTTTAAAGCTGATTTTAAAGAGCTCTCCA

1879

Db

2250

CACCATTCAGATCATAGCTGGTGTCTTTAAAGCTGATTTCT--TGAAGTTCTCA

2307

Qy

1880

GGCCCCACCAAGAGGCTGATCACACCTCAGTGGGAAGAAATAAAATTTCTTCAGGT

1939

Db

2308

GCCTTCATGGAGGATGGTCAGATCCAGTGGGGAAGAGAAATAAAATTTCTTCAGGT

2367

Qy

1940

TTTAAAA 1946

Db

2368

TTTATAA 2374

RESULT 3

BUI90334

LOCUS

AGENCY 7186901 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6001834

5' mRNA sequence.

BUI90334

VERSION

BUI90334.1 GI:22704318

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1085)

AUTHORS

NIH-MGC <http://mgc.mci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM2269 row: b column: 11
High quality sequence stop: 671.

FEATURES

source

1..1085

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6001834"

/tissue_type="hepatocellular carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_100"

/note="Organ: liver; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match

Best Local Similarity

Mismatches

Conservative

0;

Mismatches

42;

Indels

9;

Gaps

3;

Qy

264

CAGACCAGCGAAGCCATCATGAAGCTGAAT--GGCCACCAGTTGGAGAACCATGCTGAA

322

Db

1

CAGACCAGCGAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCTGAA

60

Qy

323

GGTCTCTACATCCCGGATGAGCAGATAGCAGGACCTGAGAAATGGCGCGCGAGGGG

382

Db

61

GGTCTCTACATCCCGGATGAGCAGATAGCAGGACCTGAGAAATGGCGCGCGAGGGG

120

Qy

383

CTTTGGCTCTCGGGTCTAGCGCCGCGCAGGGCTCACTGTGGCAGCGGGGCCCCAGCNA

442

Db

121

CTTTGGCTCTCGGGTCTAGCGCCGCGCAGGGCTCACTGTGGCAGCGGGGCCCCAGCNA

180

Qy

443

GCAGCAGCAAGTGGACATCCCGCTTCGGTCTCGTGGTCCACCCAGTATGTGGTGCCT

502

Db

181

GCAGCAGCAAGTGGACATCCCGCTTCGGTCTCGTGGTCCACCCAGTATGTGGTGCCT

240

Qy

503

TATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAACAGACCCAGTCCAAAGTAGA

562

Db

241

TATTGGCAAGGAGGGGGCCACCATCCGCAACATCACAACAGACCCAGTCCAAAGTAGA

300

Qy

563

CGTGATAGGAAGGAGAACGCGAGTGCAGTGAAGAGCCATCAGTGTGCATCCACCC

622

Db

301

CGTGATAGGAAGGAGAACGCGAGTGCAGTGAAGAGCCATCAGTGTGCATCCACCC

360

Qy

623

TCAGGGTCTCTCTCGCTTGAAGATGATCTTGGAGATTTATGATTAAGAGGCTTAAGA

682

Db

361

TCAGGGTCTCTCTCGCTTGAAGATGATCTTGGAGATTTATGATTAAGAGGCTTAAGA

420

Qy

683

CACCAAAACGGCTGACAGGTTCCCTGAGATCTTGGCCCATATAAATTTGTAGGGG

742

Db

421

CACCAAAACGGCTGACAGGTTCCCTGAGATCTTGGCCCATATAAATTTGTAGGGG

480

Qy

743

TCTCATTCGCAAGGAGGACGAACTGAAAGAGTGAAGAGTACCGAGACAAAT

802

Db

481

TCTCATTCGCAAGGAGGACGAACTGAAAGAGTGAAGAGTACCGAGACAAAT

540

Qy

803

CACCATCTCTCTGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGA

862

Db

541

CACCATCTCTCTGTTGCAAGACCTTACCTTTACACCTGAGAGGACCATCACTGTGA

600

Qy

863

GGGGGCGCATCGAGAAATTTGTCAGGCGCCGAGCAGGAAATAATGAAGAAAGTTCCG

922

Db

601

GGGGGCGCATCGAGAAATTTGTCAGGCGCCGAGCAGGAAATAATGAAGAAAGTTCCG

660

Qy

923

CTATGAGAAATGATGGCTGCCATGAGC-----TCTCACTGATCCCTGGCTGACCT

976

Db

661

CTATGAGAAATGATGGCTGCCATGAGC-----TCTCACTGATCCCTGGCTGACCT

720

Qy

977

GGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCGCTCCCGAGCGCT

1036

Db

721

GGCTGTGTAGGTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCGCTCCCGAGCGCT

780

Qy

1037

TACTGGGGTGTCTCCCTATAGCTTCTTTATGAGGCTCCCGAGCAGGAGATGCTGAGGT

1096

Db

781

TACTGGGGTGTCTCCCTATAGCTTCTTTATGAGGCTCCCGAGCAGGAGATGCTGAGGT

840

Qy

1097

GTTTATTCGCCGCCAGGAGTGGCGCCATCATCGCAAGAGGGCGAGCAGATCAACA

1156

Db

841

GTTTATTCGCCGCCAGGAGTGGCGCCATCATCGCAAGAGGGCGAGCAGATCAACA

900

REFERENCE AUTHORS	3 Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, U., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861
REFERENCE AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)
REFERENCE AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 2718)
JOURNAL REFERENCE AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
JOURNAL TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. Tomohiro Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Hunko Atsugi City, Kanagawa Prefecture, Japan) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/ Location/Qualifiers 1. .2718 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:B130008A04" /db_xref="MGI:2410206" /clone="B130008A04" /clone_xref="taxon:10090" /tissue_type="parthenogenote" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="9.5 days embryo" 217..1950 /note="unnamed protein product; insulin-like growth factor 2, binding protein 1 (MGD MGI:1890357) putative"
FEATURES source	1. .2718 /codon_start=1 /protein_id="BAC32119.1" /db_xref="GI:26336871" /translation="MNKLIGLNESVTPADLEKVFAEHKISYSGFVLKSGYAFVDC PDEHWAKALETFSQVLEQKLEIEHSVPKQRKQIRNIIPQLRWEVLDSLLA OYGTVENCOVNTSESTAVVNTYSNRETROAIMKNGHOLEHAKVUYIJPDEOIT QSPENGRGSGRSGRPGSPVACAPAKQPDVPLRLVPTQVGVGALIKEGATI NLKITQSKIDVHRKENAGAEKALSIVSTPEGSSACCMILEMHHKAKOTPKTDAE VPLKLAHNNFVRLIKEGRNLLKVEQDTETKITTSLODLTYLNPRTITVKGAE NCCRAQELMKVREAYENDVAAMSLIPLGLNLAAYGLFPASSAVPPSPSVTG AAPYSFMOAQEQWQVFIPAQAVGAIIGKGQHIKOLSRFASAKIAPPPTDSK VRMVLTGPPEAQFKAGRIYGLKEENFGPEKVEKLETHIRVPASAGRVTKGSK TYNELONLTAAEVVPRDQTPENDQVIKLIHGFYASQWQKIRIDILAAQVQKQHK GQSNLAQARRK"
ORIGIN	Query Match 66.8%; Score 1299.2; DB 11; Length 2718; Best Local Similarity 83.3%; Pred. No. 0; Matches 1556; Conservative 0; Mismatches 278; Indels 33; Gaps 6; QY 108 AGAGCGCGAAAAATCCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167 Db 451 AGAGTCGGAATAACAGATCCGCAATATTCACACCTCAGCTCCGATGGGAAGTACTGAT 510 QY 168 AGCTCTCGCTCAGTATGTCAGTAGAGACTGTGACCACTGTGACCACTGACACCGAGAGTGAG 227 Db 511 AGCTCTCGCTCAGTATGTCAGTAGAGACTGTGACCACTGTGACCACTGACCACTGAAAGTGAG 570 QY 228 ACGGCACTGTGTAATGTCACTTATTCACCCGGGAGCAGACCAAGCAAGCATCATCAAG 287 Db 571 ACAGCTGTGTCACTCACTTCTAACCGGAGCAGACCAAGCAAGCATCATCAAG 630 QY 288 CTGAATGGCCACCACTGTGGAGAACCAATGCCCTGAAGGTCTCTACATCCCCGATGAGCAG 347 Db 631 CTAAATGGCCATCAACTGGAGAACCACTGCCCTGAAGGTCTCTACATCACTGATGAGCAG 690 QY 348 ATAGCACAGGACCTCAGAAATGGCCGCGAGGGGGCTTTGGCTCTCGGGGTGAGGCCCGC 407 Db 691 ATACCAAGGTCTCTGAGAAATGGCGCTCTGGAGGGCTTTGGGGTCTCGGGGCGAGCCCGG 750 QY 408 CAGGGCTCACTGTGGCAGCGGGGGCCCGCAGCAAGCAGCAGCAAGTGACATCCCCCTT 467 Db 751 CAAAGGTCTCCCGTGGCAGCAGGGGTCCAGCAAGCAGCAGCAGTGGACATCCCTCTC 810 QY 468 CGCTCTCTGGTCCCAACCCAGTATGTGGTGCCATTTATTTGGCAAGAGGGGGCCACATC 527 Db 811 CGCTCTCTGGTCCCAACCCAGTATGTGGTGCCATTTATTTGGCAAGAGGGGGCCACATC 870 QY 528 CGCAACATCAAAAACAGACCCAGTCCAAAGATAGCTGATAGGAGAGGAGAGCAGGT 587 Db 871 CGAACATCAAAAACAGACCCAGTCCAAAGATAGCTGATAGGAGAGGAGAGTGGCGG 930 QY 588 GCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCCTGCTTAAG 647 Db 931 GCTGCGGAGAGGCCCATCAGCGTGCATCAACCCCTGAAAGGTGCTCTCCCTCGCGTCAAG 990 QY 648 ATGATCTTGGAGATATGATTAAGAGGCTTAGGACACCAAAACGGCTGACGAGGTTC 707 Db 991 ATGATCTTGGAGATATGATTAAGAGGCTTAGGACACCAAAACGGCTGACGAGGTTC 1050 QY 708 CTGAAGATCTCGCCCATATAACTTTGTAGGGCGTCTCATTTGGCAAGGAGAGCAGGAAC 767 Db 1051 CTGAAGATCTCGCTCATTAACAACTTCGTCGGCGGAGCTCATTTGGCAAGGAGAGGCGGAAC 1110 QY 768 CTGAAGAGGTAGACCAAGATCCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTT 827 Db 1111 CTGAAGAGGTGAGCAGCAGACAGAGACGAAGATCAACCATCTCATCTGCTCCAGGACCTC 1170 QY 828 ACCCTTCAACCCCTGAGAGGACCATCACTGTAGGGGGGCCATCGAGAAATTTGTCAGG 887 Db 1171 AGCTCTATAACCCCTGAGAGGACCATCACTGTAGGGGGGCCATTTGAGAACTGTTTCAGG 1230 QY 888 GCCGAGCAGGAAATATGAAGAAAGTTCCGGAGGCGCTATGAGAAATGATGTGCTGCCATG 947

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 06:21:04 ; Search time 4863.33 Seconds
(without alignments)
11948.985 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtagcggaggctggg.....atttccttcagggtttaaaa 1946

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1299.2	66.8	2718	11 AK044850	AK044850 Mus muscu
2	1299.2	66.8	2780	11 AK013940	AK013940 Mus muscu
3	969.4	49.8	1085	13 BUL90334	BUL90334 AGENCOURT
4	921.4	47.3	1042	12 EM928196	EM928196 AGENCOURT

5	828.4	42.6	1042	12	BM561057	
6	745.4	38.3	950	13	BQ648383	
7	723.4	37.2	4100	11	AK088465	
c	8	722.6	919	13	BX327672	
	9	721.8	37.1	2202	11	AK011689
c	10	678.6	34.9	953	13	BQ647360
	11	671.8	34.5	831	12	BG115319
c	12	670.4	34.5	1201	13	BX401619
	13	669	34.4	712	13	BQ225582
c	14	662.2	34.0	955	13	BQ651665
	15	662	34.0	896	10	BF984962
c	16	660.2	33.9	891	13	BX372273
	17	657	33.8	764	12	BM006191
c	18	636.4	32.7	983	13	BQ643920
	19	602.2	30.9	1248	29	AY417413
c	20	590.4	30.3	982	13	BQ652180
	21	588.8	30.3	929	13	BQ652586
c	22	588.4	30.2	978	12	BG115593
	23	582.2	29.9	1173	12	BM467135
c	24	576	29.6	3592	11	AK049196
	25	575	29.5	2964	11	AK044984
c	26	557.4	28.6	1321	29	AY417411
	27	554.6	28.5	3262	11	AK077404
c	28	547.6	28.1	884	13	BU598549
	29	546	28.1	568	14	CB146278
c	30	542.6	27.9	1061	13	BQ647561
	31	540	27.7	545	13	BX089485
c	32	513.6	26.4	658	13	BQ256493
	33	513.4	26.4	656	10	AW303505
c	34	513.4	26.4	1321	29	AY417412
	35	512.4	26.3	550	9	AA978341
c	36	505	26.0	697	13	BU400982
	37	498.4	25.6	643	14	CB894671
c	38	497.2	25.5	517	10	BE550506
	39	495.4	25.5	692	13	BU460238
c	40	491.2	25.2	677	14	CB841164
	41	487.4	25.0	896	13	BQ647425
c	42	485.4	24.9	761	9	AW003366
	43	475	24.4	714	13	BU268759
c	44	469	24.1	484	9	AI337147
	45	459.4	23.6	486	9	AL704124

ALIGNMENTS

RESULT 1
AK044850
LOCUS
DEFINITION
AK044850 2718 bp mRNA linear HTC 20-SEP-2003
Mus musculus 9.5 days embryo Parthenogenote cDNA, RIKEN full-length
enriched library, clone:B130008A04 product:insulin-like growth
factor 2, binding protein 1, full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK044850
AK044850.1 GI:26336870
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
1
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

This Page Blank (uspto)

830	Db	 TCCCTTGAAGATTTTAGCTCATAATAACTTTTGTGGACGCTCTATTGTGTAAGAAGAA	889
763	Qy	GGAACTGAAGAAGGTAGACCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCGAAG	822
890	Db	GAATCTTAAAAAAATTTGACACAGACACAGACTAAATCAGATATCTCATTTGCAGG	949
823	Qy	ACCTTACCCCTTTACACCCCTGAGAGGACCATCTGTGAAGGGGGCCATCGAGAAATTGTT	882
950	Db	AATTGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG	1009
883	Qy	GCAGGGCCGAGCAGGAAAATTAATGAAGAAATTCGGAGCGCCTATGAGATGATGTGGCTG	942
1010	Db	CCAAAGCTGAGGAGGAGATCATGAAGAAAATCAGGAGTCTTATGAAAATGATATTGCTT	1069
943	Qy	CCATGA-----GCTCTCACCTGATCCCTGGCCTGAACTGGCTGCTGTAGGTCTTTTCC	996
1070	Db	CTATGAATCTTCAAGCACTTTAATCTCTGGATTAATCTGAACGCTTGGGTCTGTTCC	1129
997	Qy	CAGCTTCATCCAGCGCAGTCCCGCGCTCCACAGACGCTTACTGGGGTGCTCCCTATA	1056
1130	Db	CACCACTTCAGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCATGACTCTCCCT	1189
1057	Qy	GCTCCTTTATGCAGGCTCCGAGCAGAGATGTCAGGTGTTTATCCCGCCCGCAGGAG	1116
1190	Db	ACCGCGAGTTTGAGCAATCAAGAACGGAGACTGTTTCATCTGTTTATCCAGCTCTATCAG	1249
1117	Qy	TGGCGCGCATCATCGGCAAGAAGGGCAGACATCAAAACAGCTCTCCCGGTTTGGCAGGG	1176
1250	Db	TCGGTGCATCATCGGCAAGCAGGGCCAGCACATCAGACAGCTTCTCGCTTTGCTGGAG	1309
1177	Qy	CCTCCATCAAGATTGCACCAACCCGAAACACTGACTCCAAAGTTGCTATGTTTATCATCA	1236
1310	Db	CTTCAATTAGATTGCTCCAGCGGAAGCACCAAGTCTAAAGTGAGGATGGTGATTATCA	1369
1237	Qy	CTGACCGCCAGAGGCCCAATTCAAGCTCAGGGAGAACTATGCGAACTCAAGGAGS	1296
1370	Db	CTGACCACCAAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTATGAAAAAATTAAGAAG	1429
1297	Qy	AGAACTTCTTTGTTGCCAAGGAGGAAGTGAAGCTGGAGACCAATACGTGTGCCAGCAT	1356
1430	Db	AAAACTTTGTTAGTCTTAAGAGAGGTGAAACTTTGAGCTCATATCAGAGTGCCATCCT	1489
1357	Qy	CAGCAGTGGCCGGTCAATTGGCAAAAGGTGGAAAAACGGTGAAACGAGTTCGAGAAATTGA	1416
1490	Db	TTGCTGCTGGCAGAGTTATTGGAAAAAGGCAAAACGGTGAATGAACTTTCAGAAATTTGT	1549
1417	Qy	CGGCAGCTGAGGTGGTAGTACCAGAGACAGACCCCTGATGAGACGACCGCATCG	1476
1550	Db	CAAGTCGAGAAGTTGTTGTCCCTCGTGACACAGACCTGATGAGATGACCAAGTGGTTG	1609
1477	Qy	TGAAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACCGAAGATCCGAGACATCC	1536
1610	Db	TCAAAATAACTGGTCACTTCTATGCTTGGCAGGTTGCCAGAGAAAAATTCAGAGAAATTC	1669
1537	Qy	TGGCCCGAGTTAAGCAGCAGCATCAGAAG	1565
1670	Db	TGACTCAGGTTAAGCAGCACCAACAACAG	1698

Search completed: July 24, 2004, 22:28:55
Job time : 843.722 secs

523 CCATCCGCAACATCACAAAAAGCCAGTCCAGATAGAGCTGATAGGAGGAGACG 582
|||
Db |||
583 CAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCCTCGGCTT 642
|||
Db |||
710 CGGGGGCTGCTGAGAGTCGATTACTCTCTCTACTCTCTGAGGACCTCTCGGGCTT 769
|||
Qy |||
643 GTAAGATGATCTTGAGATTTATGATTAAGAGGCTTAAGGACCAAAAAAGGCTGACGAGG 702
|||
Db |||
770 GTAAGTCTTATCTGGAGATTATGATTAAGGAGCTTAAGGACCAAAAAAGGCTGACGAGG 829
|||
Qy |||
703 TTCCCTCTGAAGATCTCGCCCATATTAATCTTGTAGGGCGTCTCATTTGGCAGGAGGAC 762
|||
Db |||
830 TCCCTCTGAAGATTTTATGATTAATTAATCTTGTAGGCGTCTTATTTGGTAAAGAGAA 889
|||
Qy |||
763 GGAACCTGAAGAAGGTAGAGCAAGATACCGAGACAAAAATCACATCTCTCTGTTGCAAG 822
|||
Db |||
890 GAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAAAATCACGATATCTCCATTGCGG 949
|||
Qy |||
823 ACCTTACCTTTTACAACTGAGAGGACCATCTCTGTAGGGGGCCATCGAGAAATGTTT 882
|||
Db |||
950 AATTGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATTTTGAGACATGTG 1009
|||
Qy |||
883 GCAGGGCGGAGCAGGAAATTAATGAAGAAAGTTTCGGGAGGCTATGAGAAATGATGTGCTG 942
|||
Db |||
1010 CCAAGCTGAGGAGGAGATCATGAGAAATCAGGGAGTCTTATGAAATGATATGCTT 1069
|||
Qy |||
943 CATTGA-----GCTCTCACTGATCCCTGCTGAACTGCTGTGATGCTTCTTCC 996
|||
Db |||
1070 CTATGATCTTCAAGCACATTTAATCTCTGATTAATCTGAAGCCCTTGGGCTGTTCC 1129
|||
Qy |||
997 CAGCTTCATCAGCGAGTCCGCGGCTCCAGCAGCGTACTGGGGCTGCTCCCTATA 1056
|||
Db |||
1130 CACCCACTTCAGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGACTCTCCCT 1189
|||
Qy |||
1057 GCTCTCTTATGCAAGCTCCGAGCAGGAGATGTTGCAAGTGTATTCCTCCGCCCCAGGAG 1116
|||
Db |||
1190 ACCCGAGTTTGAACATCAGAAACGAGACTGTTATCTGTTTATCCAGCTTATCAG 1249
|||
Qy |||
1117 TGGGGCGCATCATCGGCAAGAGGGGAGCAGCATCAAAAGCTCTCCCGTGTGCCAGCG 1176
|||
Db |||
1250 TCGGTGCAATCATCGGCAAGCAGGGCCAGCACATCAAGCAGCTTCTCGCTTTGCTGGAG 1309
|||
Qy |||
1177 CTTCCATCAAGATTGCAACACCGAAACACTGACTCCAAAGTTCGTATGTTATCATCA 1236
|||
Db |||
1310 CTTCAATTAAGATTGCTCCAGCGAAGCACAGATGCTAAAGTGAGGATGTTGATTATCA 1369
|||
Qy |||
1237 CTGGACCGCAGAGGCCCAATTCAGAGCTCAGGGAAGAACTTATGGCAAACTCAAGGAGG 1296
|||
Db |||
1370 CTGGACCAACAGAGGCTCAGTTCBAGGCTCAGGGAAGATTTATGAAATTAAGAAG 1429
|||
Qy |||
1297 AGAATCTTTGTTCCCAAGAGGAAGTGAAGCTGGAGACCCACATAGTGTGCCAGCAT 1356
|||
Db |||
1430 AAAAATTTGTTAGTCTTAAAGAAAGAGTGAACTTGAAGCTCATATCAGAGTGCCATCT 1489
|||
Qy |||
1357 CAGCAGTGGCGGGTCAATTGGCAAGGTGGAAGAAAGGTGAACGAGTTCAGAAATTTGA 1416
|||
Db |||
1490 TTGCTGCTGCAGAGTTATTGGAAGAGGAGGCAAAAGCGTGAATGAATTCAGAAATTTGT 1549
|||
Qy |||
1417 CGGACGTGAGGTGGTATGATACCAAGAGACAGACCCCTGATGAGAACGACAGGTATCG 1476
|||
Db |||
1550 CAAAGTCAGAAGTTGTTGTCCTCGTGACCAAGACACTGATGAGATGACCAAGTGGTTG 1609
|||
Qy |||
1477 TGAATATCATCGACATTTCTATGCCAGTCAGATGGCTCAACGGAGATCCGAGACATCC 1536
|||
Db |||
1610 TCAAAATAAATGTTGCTATCTATGCTTCCAGGTTGCCAGAGAAAAATTCAGGAAATTC 1669
|||
Qy |||
1537 TGGCCAGGTTAAGCAGCAGCATCAGNAG 1565
|||
Db |||
1670 TGAATCAGGTTAAGCAGCAGCAACAAACAG 1698
|||

RESULT 15

US-09-897-778-447
; Sequence 447, Application US/09897778

; Patent No. US20020147143A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Peckham, David W.

; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; CURRENT APPLICATION NUMBER: US/09/897,778

; CURRENT FILING DATE: 2001-06-28

; NUMBER OF SEQ ID NOS: 467

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 447

; LENGTH: 1743

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-897-778-447

Query Match 34.6%; Score 674.2; DB 9; Length 1743;

Best Local Similarity 66.7%; Pred. No. 6.6e-200;

Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

Qy 103 GGGTTAGGAGCCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTAC 162
Db |||
Qy 230 GGCAAGGATTCGAAACTTCAGATACGAAATATCCGCTCTATTACAGTGGAGTGC 289
|||
Db |||
Qy 163 TGCACAGCTGCTGGCTCAGTATGTTACAGTACAGAACTGTGAGCAAGTGAACCCGAGA 222
|||
Db |||
Qy 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGAGAGCTGTGAGCAAGTGAACCTGACT 349
|||
Qy 223 GTGAGACGGCAGTGTGAATGTCACTTATCCAAACCGGAGCAGACCCAGCAAGCCATCA 282
|||
Db |||
Qy 350 CGGAAACTGCAGTTGTAATGTAACTTATCCAGTAAAGCAGCAAGCTAGCAAGCTAG 409
|||
Qy 283 TGAAGCTGAATGSCCAGCTGTGGAGAACCATGCTCCCTGAAGTCTCTCATCCCCGATG 342
|||
Db |||
Qy 410 ACAAACTGAATGGATTTTCAGTTAGAGAAATTTCACTTGAAGTAGCTATATCCCTGATG 459
|||
Qy 343 AGCAGATAGCACAGGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTGAGC 402
|||
Db |||
Qy 470 AAACGGCCGCGCAAGAAACCCCTTGCAGCAGCCCGAGGTGCGCCGGGGCTTTGGGCGAG 529
|||
Qy 403 CCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAGCAGCAGCAAGTGGACATCC 462
|||
Db |||
Qy 530 GGGGCTCCCAAGGCAAGGCTCTCCAGGATCCGTATCCAGCAGAACCATGTGATTGTC 589
|||
Qy 463 CCCTTCGGCTCTGTTGGTCCCAAGCAGTATGTTGGTGGCTTATTTGTCAGAGGGGGGCGCA 522
|||
Db |||
Qy 590 CTCTGGCTGCTGTTGCTCCCAAGCAATTTGTTGGAGCCATCATAGGAAAGAGGTGCCA 649
|||
Qy 523 CCATCCGCAACATCACAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTCTCTCCGCTT 582
|||
Db |||
Qy 650 CCATTCGGAACATCACAAACAGACCCAGTCTAAATTCGATGTCCACCTTAAGAAAAATG 709
|||
Qy 583 CAGGTSAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTCTCTCCGCTT 642
|||
Db |||
Qy 710 CGGGGCTCTCTGAGAAGTCTGATTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
|||
Qy 643 GTAAGATGATCTTGAGATTTATGATTAAGAGGCTTAAGGACCAAAAAAGGCTGACGAGG 702
|||
Db |||
Qy 770 GTAAGTCTTATCTGGAGATTATGATTAAGGAGCTTAAGGACCAAAAAATTCAGAGAGAG 829
|||
Qy 703 TTCCCTGAAGATCTTGGCCCATATAACTTTTGGGGCTCTCATTTGCCAAGGAGGAC 762
|||

1430 AAACTTTGTTAGTCTTAAAGAGAGGTTGAAACTTGAAGCTCATATCAGAGTGCATCT 1489
1357 CAGCAGCTGGCGGCTTATTGGCAAGGTGGAAGAAACGGTGAACGAGTTGCAGAAATTGA 1416
1490 TTGCTGTGGCAGAGTTTATTGGAAGAGGAGCNAACGGTGAATGAATCTTCAAGAAATTGT 1549
1417 CGGCAGCTGAGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACACGAGTTCATCG 1476
1550 CAAGTGCAGAAAGTTGTTGCTCCCTGTGACCAAGACACCTGATGAGNAATGCCAAGTGGTTG 1609
1477 TGAATAATCATCGACATTTTATGTCAGTGTGATGCTCAACGGAGAGTTCGAGACATCC 1536
1610 TCAATAATACCTGTCTTCTTCTGTCAGGTTGCCAGAGAAAAATTCAGGAAATTC 1669
1537 TGGCCAGGTTAAGCAGCAGCATCAGAG 1565
1670 TGACTCAGGTAAAGCAGCACCACCAACAG 1698

RESULT 14
US-10-313-986-478
; Sequence 478, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-478

Query Match 34.6%; Score 674.2; DB 16; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

103 GGGTTAGGAGCGCGGAAATTCAAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTAC 162
230 GGCAGAGGATTCCGAAACTTCAGATACGAAATATCCCGCTCATTTACAGTGGGAGTGC 289
163 TGGACAGCTGCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAGTGAACACCGAGA 222
290 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
223 GTGAGACGCGAGTGGTGAATGTCACTATTCAAACGGAGACGACGAGCAAGCCATCA 282
350 CGGAACCTGCATTTGAATGTAACCTATCCAGTAGAGCCAGCTAGCAGCACTAG 409
283 TGAAGCTGAATGGCCACCAAGTTGGAGAACCAATGCCCCGAGGTTCTTACATCCCCGATG 342
410 ACAAACTGAATGGATTTCAGTTAGAAATTTACCTTTGAAAGTAGCTATATCCCTGATG 469
343 ACAGATAGCAGAGGACCTGAGATGGGGCGGAGGGGCTTTGGCTCTCGGGGTCAGC 402
470 AAAACGCGCGCCAGCAAAAAACCCCTTTCAGACGCCCCGAGGTGCGCGGGGCTTGGGCGA 529
403 CCGCCAGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAGTGAACATCC 462
530 GGGGCTTCCTCAGGCGAGGGGTCTCCAGGATCCGATCCAGCAGAAACCATGTGATTGC 589
463 CCCTTCGGCTCCTGTGGTCCCAACCATGATGTTGGTGCATTTGGCAAGAGGGGGCCA 522
590 CTCCTGGCTGCTGGTTCACCCCAATTTGTTGGAGCCATCATAGGAAGAGAGTGC 649
523 CCATCCGCAACATCAAAAACAGACCCAGTCCAGATAGACGTGCATAGGAAGAGAGACG 582
650 CCAATTCGGAACATCACCAACAGACCCAGTCTAATAATCGATGTCACCGTTAAAGAAATG 709
583 CAGGTGACGTGAAAGGCAATCAGTGTGCACTCCACCCCTGAGGCTGCTCTCCGCTT 642
710 CGGGGCTGCTGAGAAAGTCGATTACTTCTCTCTACTCTCTGAGGCACTCTGCGGCTT 769
643 GTAAGATCATCTTCGAGATTATGCATAAGAGGCTAAGACACCAAAACGGCTGACGAGG 702
770 GTAAGTCTATCTCGAGATTATGCATAAGAGGCTAAGATATAAATTCAGAGAGAGA 829
703 TTCCCTCAGAGATCCTGCGCCCAATAATCTTTGTAGGGGCTCTCATTTGSCAAGAGAGAC 762
830 TCCCTTCGAGATTCTAGCTCATATAATCTTTGTTGGAGCTCTTATTTGTAAGAGAGAA 889
763 GGAACCTGAAGAGGTGAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAG 822
890 GAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAAATCAGATATCTCATTTGCAAG 949
823 ACCTTACCTTTACAACTGAGAGACCATCACTGTGAAGGGGCGCATCGAGAAATGTT 882
950 AATTGACGCTGTATTAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009
883 GCAGGCGCGAGAGAGAAATTAATCAAGAAAGTTTCGAGAGGCTTATGAGAAATGATGGCTG 942
1010 CCAAGCTGAGAGAGAGATCATGAGAAATCAGGAGTCTTATGAAATGATATGCTT 1069
943 CAATGA-----GCTCTACCTGATCCCTGGCTGAACTGCTGCTAGTCTTTTCC 996
1070 CTATGAATCTTCAAGCACATTTAATCTCTGGATTAAATCTGAACGCTTTGGGCTGTGTTCC 1129
997 CAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTATA 1056
1130 CACCACTTCAGGATGCCACCTCCACCTCAGGCGCCCTTCAGCCATGATCTCTCCCT 1189
1057 GCTCTCTTATGAGGCTCCCGAGCAGGAGATGGTGCAGTGTTTATPCCCGCCCGAGCAG 1116
1190 ACCGCGAGTTTGAGCAATCAGAAACGAGACTGTTCTATCTGTTATCCAGCTCTATCAG 1249
1117 TGGGCGCCATCATCGCAGAGAGGGGAGCAGCATCAACAGCTCTCCCGTTTGGCAGCG 1176
1250 TCGGTGCCATCATCGCAGAGAGGGGCGCAGCATCAAGCAGCTTTCTCTGCTTGTGGAG 1309
1177 CTTCCATCAAGATTGCAACACCGAAACCTTGACTCCAAAGTTTCGTATGTTTATCATCA 1236
1310 CTTCAATTAAGATTGTTCCAGCGAAGACCAAGATGCTAAGTGAGATGGTATATCA 1369
1237 CTGGACCGCCAGAGGCGCAATTCAGGCTCAGGAGAAATCTATGAGCAAACTCAAGAGG 1296
1370 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGAGAAATTTATGAAATAATTAAGAGAG 1429
1297 AGAATCTTTTGTCCCAAGGAGAGTGAAGTGGAGAGCCCAATACGTGTGTCAGCAT 1356

350 CGGAAACTGCAGTTGTAATGTAACTATTCCAGTAAAGGACCAAGCTAGACAAGCACTAG 409
283 TGAAGCTGAATGGCCACAGTTGGAGAACCATGCTTGAAGGTCTCTTACATCCCGGATG 342
410 ACAAACTGAATGGATTTCAGTTAGAGAAATTTCACTTGAAGTAGCTTATCCCTGATG 469
343 AGCAGATAGCAGAGGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGTTCAGC 402
470 AAAACGCGCGCCAGCAAAACCCCTTGCAGCAGCGCCCGAGGTGCGCGGGGCTTGGGCGA 529
403 CCGCCAGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCC 462
530 GGGGCTCTCTCAAGGCGAGGGTCTCCAGGATCCGATCCCAAGCAGAAACCATGTGATTTC 589
463 CCCTTCGGCTCCTGTGTCGCCACCATGATGTTGGTGCATTTGGCAAGAGGGGGCCA 522
590 CTCCTGGCTGCTGGTTCACCCCAATTTGTTGGAGCCATCATAGGAAGAGAGTGC 649
523 CCATCCGCAACATCAAAAACAGACCCAGTCCAGATAGACGTGCATAGGAAGAGAGACG 582
650 CCAATTCGGAACATCACCAACAGACCCAGTCTAATAATCGATGTCACCGTTAAAGAAATG 709
583 CAGGTGACGTGAAAGGCAATCAGTGTGCACTCCACCCCTGAGGCTGCTCTCCGCTT 642
710 CGGGGCTGCTGAGAAAGTCGATTACTTCTCTCTACTCTCTGAGGCACTCTGCGGCTT 769
643 GTAAGATCATCTTCGAGATTATGCATAAGAGGCTAAGACACCAAAACGGCTGACGAGG 702
770 GTAAGTCTATCTCGAGATTATGCATAAGAGGCTAAGATATAAATTCAGAGAGAGA 829
703 TTCCCTCAGAGATCCTGCGCCCAATAATCTTTGTAGGGGCTCTCATTTGSCAAGAGAGAC 762
830 TCCCTTCGAGATTCTAGCTCATATAATCTTTGTTGGAGCTCTTATTTGTAAGAGAGAA 889
763 GGAACCTGAAGAGGTGAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAG 822
890 GAAATCTTAAAAAATTTGAGCAAGACACAGACACTAAAATCAGATATCTCATTTGCAAG 949
823 ACCTTACCTTTACAACTGAGAGACCATCACTGTGAAGGGGCGCATCGAGAAATGTT 882
950 AATTGACGCTGTATTAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009
883 GCAGGCGCGAGAGAGAAATTAATCAAGAAAGTTTCGAGAGGCTTATGAGAAATGATGGCTG 942
1010 CCAAGCTGAGAGAGAGATCATGAGAAATCAGGAGTCTTATGAAATGATATGCTT 1069
943 CAATGA-----GCTCTACCTGATCCCTGGCTGAACTGCTGCTAGTCTTTTCC 996
1070 CTATGAATCTTCAAGCACATTTAATCTCTGGATTAAATCTGAACGCTTTGGGCTGTGTTCC 1129
997 CAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTATA 1056
1130 CACCACTTCAGGATGCCACCTCCACCTCAGGCGCCCTTCAGCCATGATCTCTCCCT 1189
1057 GCTCTCTTATGAGGCTCCCGAGCAGGAGATGGTGCAGTGTTTATPCCCGCCCGAGCAG 1116
1190 ACCGCGAGTTTGAGCAATCAGAAACGAGACTGTTCTATCTGTTATCCAGCTCTATCAG 1249
1117 TGGGCGCCATCATCGCAGAGAGGGGAGCAGCATCAACAGCTCTCCCGTTTGGCAGCG 1176
1250 TCGGTGCCATCATCGCAGAGAGGGGCGCAGCATCAAGCAGCTTTCTCTGCTTGTGGAG 1309
1177 CTTCCATCAAGATTGCAACACCGAAACCTTGACTCCAAAGTTTCGTATGTTTATCATCA 1236
1310 CTTCAATTAAGATTGTTCCAGCGAAGACCAAGATGCTAAGTGAGATGGTATATCA 1369
1237 CTGGACCGCCAGAGGCGCAATTCAGGCTCAGGAGAAATCTATGAGCAAACTCAAGAGG 1296
1370 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGAGAAATTTATGAAATAATTAAGAGAG 1429
1297 AGAATCTTTTGTCCCAAGGAGAGTGAAGTGGAGAGCCCAATACGTGTGTCAGCAT 1356


```
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-478

Query Match      34.6%; Score 674.2; DB 15; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGGACCGGAAATTCAAATCCGAAATATTCACCCAGCTCCCGATGGGAAGTAC 162
D 1117 TGGGCGCCATCATCGGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGTTTGGCAGCG 1176
D 1190 ACCCGAGTTTGAGCAATCAGAAACGAGAGCTGTCATCTGTTATATCCAGCTCTATCAG 1249
QY 1117 TGGGCGCCATCATCGGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGTTTGGCAGCG 1176
D 1250 TCGGTGCGCATCATCGGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGTTTGGCAGCG 1309
QY 1177 CTTCATCAAGATTGACCAACCCGAAACCTGACTCTCAAGTTCGATGTTATCATCA 1236
D 1310 CTTCATCAAGATTGCTCCAGCGAAGCACCAGATGCTAAAGTCAGGATGTTATCATCA 1369
QY 1237 CTGACCGCCAGAGGCCAATTCAGGCTCAGGGAAGATCTATGGAACCTCAAGAGG 1296
D 1370 CTGACCGCCAGAGGCCAATTCAGGCTCAGGGAAGATCTATGGAACCTCAAGAGG 1429
QY 1297 AGAATCTTCTTGTCCCAAGGAGGAGTGAAGCTGGAGCCACATACGTCGTCAGCAT 1356
D 1430 AGAATCTTGTAGTCTTAAAGAGAGGTGAATCTGAAGCTCATATCAGAGTGCATCT 1489
QY 1357 CAGCAGCTGCGCGGTCATTGGCAAGGTTGGAAGGAGGTCGAGTTCAGAAATTTGA 1416
D 1490 TTGCTGCTGCGCAGAGTTATTGGAAAGAGGAGGCAAAAGCGTGAATGAATTCAGAA 1549
QY 1417 CGCAGCTGAGGTGATGATACCAAGAGACAGACCCCTGATGAGAACGACAGGTCATCG 1476
D 1550 CAAAGTCGAGAAAGTTGTTGTCCTCTGTCGTCGAGCAGACCTGATGAGAAATGACCA 1609
QY 1477 TGAATATCATCGGACATTTCTATGCGCAGTCAGATGCTCAACGGAAGATCCGAGACATCC 1536
D 1610 TCAATATTAATCTGTCATCTTCTGTCGAGGTTGCCAGAGGTTCCAGAGAAATTCAG 1669
QY 1537 TGGCCAGGTTTAAAGCAGCAGCATCAGAAG 1565
D 1670 TGAATCAGGTAAAGCAGCAGCAGCAACAG 1698

RESULT 13
US-10-313-986-347
; Sequence 347, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-313-986-347

Query Match      34.6%; Score 674.2; DB 16; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGGACCGGAAATTCAAATCCGAAATATTCACCCAGCTCCCGATGGGAAGTAC 162
D 230 GGCAGAGGATTCGAAACTTCAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGC 289
QY 163 TGGACAGCTGCTGGCTCAGTATGGTATCAGTACAGAACTGTGAGCAAGTGAACACGAGA 222
D 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTCCAAACGGGAGCAGACCCAGGCAAGCCATCA 282
D 350 CGGAACTGCAATGTTGAATGTAACTTCCAGTAAGGACCAAGCTAGACAGCACTAG 409
QY 283 TGAAGCTGAATGCGCCACCTGTTGAGAACCATGCTTGAAGGTCTCTCTACATCCCGGATG 342
D 410 ACAAACTGAATGGAATTCAGTTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG 469
QY 343 AGCAGATAGCACAGGACCTGAGNATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGC 402
D 470 AAACGGCGCCGAGCAAAACCCCTTGCAGCAGCGCCCGAGGTCGCGGGGGCTTTGGGCGA 529
QY 403 CCGCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAAGTGCACATCC 462
D 530 GGGGCTCTCAAGGCGGGGTCTCCAGGATCCGATTCAGAGCAAGAACCTATGTTGTC 589
QY 463 CCTTGGGCTCTGGTGCCCAACCAAGTATGGGTGCCATTAATTTGGCAAGGAGGGGCCA 522
D 590 CTCTGGGCTCTGGTGTCCACCACTTGTGGAGCCATCATAGGAAAGAAAGAGGTGCCA 649
QY 523 CCATCCGACATCACAACACAGACCCAGTCCAGATAGACGTGCTATGGAAGAGAGACG 582
D 650 CCATTCGGAACATCACCACACAGCCAGCTAAATTCGATGTCACCGTAAAGAAATG 709
QY 583 CAGGTGAGCTGAAAAGCCATCAGTGTGACCTCCACCCCTGAGGGTGTCTCTCGGCTT 642
D 710 CGGGGCTCTGAGAGTGCATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
QY 643 GTAAGATGATCTTGGAGATTTATGATTAAGAGGCTAAGGACACCAAAACGGCTGAGGAG 702
D 770 GTAAGTCTATTCTGGAGATTTATGATTAAGAGGCTAAGGACCTAAGATATAAAATTCAC 829
QY 703 TTCCCTGGAATCTGGCCCAATAATACTTTAGGGGCTCTCATTTGGAAGAGGAGGAC 762
D 830 TCCCTTGAAGATTTTACTGCTAATAAATCTTTGAGGCTCTTATTTGGTAAAGAGAGAA 889
QY 763 GGAACCTGAGAGGTTAGACCAAGATACCGAGACAAATTCACCATCTCTCTCTCTGCAAG 822
D 890 GAAATCTTAAATAATTGACACAGACACACACACTAAATTCAGATATCTCAATTCGAGG 949
QY 823 ACCTTACCTTTACACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAATTTGTT 882
D 950 AATTGACGCTGTATAATCCAGAACCGCACTATTACAGTTAAAGCAATGTTGAGACATGTG 1009
QY 883 GCAGGGCCGAGGAGAAATTAACAGAAAGTTCCGGAGGCTCATGAGATGATGGCTG 942
D 1010 CCAAGCTGAGGAGGAGATCATGAGAAATTCAGGGAGTCTTATGAAATGATATTGCTT 1069
QY 943 CCATGA-----GCTCTCACTGATCCCTGGCTGAACTGGCTGCTGTAGTGTCTTTCC 996
D 1070 CTATGATCTTCAAGCACATTTAATCTCTGATTAATCTGAAGCCCTTGGGCTGTGTTCC 1129
QY 997 CAGTTCATCCAGGACAGTCCCGCGGCTCCAGAGCGGTTACTGGGGCTGCTCCCTATA 1056
D 1130 CACCCATTCAGGAGTGCACCTCCCACTCAGGGGCCCCCTTCAGGCCATGACTCTCCCT 1189
```

APPLICANT: Mericle, Barbara
APPLICANT: Spies, Gregory A.
APPLICANT: Fan, Liqun
APPLICANT: Wang, Tongtong
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C18
CURRENT APPLICATION NUMBER: US/10/117,982
CURRENT FILING DATE: 2002-04-05
NUMBER OF SEQ ID NOS: 484
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 347
LENGTH: 1740
TYPE: DNA
ORGANISM: Homo sapiens
US-10-117-982-347

Query Match 34.6%; Score 674.2; DB 15; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
QY 103 GGGTTAGAGCGCGAAATTCAGAAATATTCACCCAGCTCCGATGGGAAGTAC 162
DB 230 GGCAGAGATTCGGAATCTCAGATACGAAATATCCGCCCTCATTTACAGTGGAGGTGC 289
QY 163 TGGACAGCTCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCGAGA 222
DB 290 TGGATAGTTTACTAGTCAGTATGGAGTGGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
QY 223 GTGACAGCGAGTGGTGAATGTACCTATTCACCCGGGAGGAGCAGCAGGCAAGCCATCA 282
DB 350 CGGAACTCGAGTTGTAATGTAACTTACCTATTCAGTAAAGGACCAAGCTAGAACACTAG 409
QY 283 TGAAGCTGAAGTGGCCAGCTGGAGAACCATGCTCCAGAGTCTCCATACATCCCGCATG 342
DB 410 ACAAACTGAATGGAATTCAGTTAGAGAAATTCACCTTGAAAGTAGCCTATATCCCTGATG 469
QY 343 AGCAGATAGCAGGAGCCTGAGAAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTACGC 402
DB 470 AAACGGCGCCAGCAAAACCCCTTGACAGAGCCCGAGGTGCGGGGGCTTGGGCGAGA 529
QY 403 CCGGCGAGGGTCACTGTGGAGCGGGGGCCAGCCAGCAGCAGCAGCAGTGGACATCC 462
DB 530 GGGGCTCTCAAGGCGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATGTGATTTGC 589
QY 463 CCCTTCGGCTCTGCTGCGCCAGTATGGGTGCGCATTTTGGCAAGGAGGGGCCA 522
DB 590 CTCTGCGCTCTGCTGCTCCACCCCAATTTGTTGGAGCCATCATAGAAAGAGGTGCCA 649
QY 523 CCATCCGCAACATCAAAAAACAGCCAGTCCAGATAGACCTGTCATAGGAAGGAGAAG 582
DB 650 CCATTCGGAAATCAACCAACAGACCCAGTCTAAATCGATGCCACCGTAAAGAAATG 709
QY 583 CAGTGCAGCTGAAAGGCAATCAGTGTGCACTCCACCCCTGAGGGCTCTCTCGGCTT 642
DB 710 CGGGGGCTGCTGAGAAGTCGATTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
QY 643 GTAAGTATCTTTCAGAGATTTATGATAAGAGGCTTAAGGACACCAAAACGGCTGACGAG 702
DB 770 GTAAGTCTATCTTGAGATTTATGATAGGAAGCTCAAGATATAAAATTCAGAGAGA 829
QY 703 TTCCCTCTGAAGATCTCGGCCCAATAATACTTTGTAGGGCTCTCATTTGGCAAGGAGGAC 762
DB 830 TCCCTTGAAGATTTTAGCTCATATAACTTTTGGAGCTCTTATTTGTAAGAGAGGAA 889
QY 763 GGAACCTGAGAGGTAGGAGATACCGAGACAAATACCATCTCTCTCTCTCTCTCTCTCT 822
DB 890 GAAATCTTAAAGAAATTTAGCAGAGACACAGACACTAAATACAGATATCTCCATTCAGG 949
QY 823 ACCTTACCTTTTACAACTCTGAGAGACCATCACTGTGAAAGGGGGCCATCGAGAAATGTT 882
DB 950 AATTGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGCAATGTTGAGACATGTG 1009

QY 883 GCAGGCGGAGCAGGAATATATGAAGAAAGTTCCGGAGGCTTATGAGAAATGATGTGGCTG 942
DB 1010 CCAAGAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGATATGTGTT 1069
QY 943 CCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTGTAGTCTTTTCC 996
DB 1070 CTATGAATCTCAAGCACAATTAATTCCTGATTAAATCTGAACGCTTGGGTCTGTTC 1129
QY 997 CAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGGTTACTTGGGGCTGCTCCCTATA 1056
DB 1130 CACCACTTCAGGGATGCCACTCCACCTCAGGCGCCCTTCAGCCATGACTCTCCCT 1189
QY 1057 GCTCCTTTATGACGCTCCCGAGCAGGAGATGTGAGTGTATTATCCCGCCAGGAG 1116
DB 1190 ACCCGAGTTTGAAGATCAGAAACGGAGACTGTTTCATCTGTTATCCAGCTCTATCAG 1249
QY 1117 TGGCGCCCATCATCGCAAGAGGGGAGCAGCACAACACAGCTCTCCCGTTTCCAGCG 1176
DB 1250 TCGGTGCCATCATCGGAGCAGGCGCAGCAGCACAACAGCAGCTTCTCGCTTGGAG 1309
QY 1177 CCTCATCAAGATTGCACCCAGAACACCTGACTCCAAAGTTCGATGTTGTTATCATCA 1236
DB 1310 CTTCAATTAAGATTGCTCCAGCGGAAGCAGAGATGCTAAAGTGAAGTGGTATATCA 1369
QY 1237 CTGACCCCGCAGAGCCCAATTCAGGCTCAGGGAAGATCTATGSCAAACTCAAGGAGG 1296
DB 1370 CTGACCAACCCAGAGCTCAGTTCAGGCTCAGGGAAGATTTATGGAAGAAATTAAGAG 1429
QY 1297 AGAACTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCGAGCAT 1356
DB 1430 AAACCTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGTCTATATCAGAGTGCATCT 1489
QY 1357 CAGCAGCTGCGCGGTCTATTTGGCAAGGTGGAAGAAACCGTGAACAGTTCGAGAATTGA 1416
DB 1490 TTGCTGCTGGCAGATTATTTGAAAAAGGAGGCAAAACCGTGAATGAACCTTCAGAAATTTG 1549
QY 1417 CGGACGTGAGGTGTAGTACCAAGACACGACCCCTGATGAGAACGACCAAGTGCATCG 1476
DB 1550 CAAGTGCAGAAAGTTGTTGTCCTGTCGACGACACCTGATGAGATGACCAAGTGGTG 1609
QY 1477 TGAATCATCGACATTTCTATGCGAGTTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
DB 1610 TCAATTAATCTGTCATCTTCTATGCTTGCAGTTTCCCGAGAGAAAAATTCAGGAAATTC 1669
QY 1537 TGGCCCGAGTTAAGCAGCAGCATCAAG 1565
DB 1670 TGACTCAGTAAAGCAGCACCAACACAG 1698

RESULT 12
US-10-117-982-478
; Sequence 478, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 478

APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Li, Samuel X.
APPLICANT: Kalos, Michael D.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
APPLICANT: Retter, Marc W.
APPLICANT: Durham, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Vedwick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Peckman, David W.
APPLICANT: Cai, Feng
APPLICANT: Foy, Teresa M.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C17
CURRENT APPLICATION NUMBER: US/10/007,700
NUMBER OF SEQ ID NOS: 469
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 347
LENGTH: 1740
TYPE: DNA
ORGANISM: Homo sapiens
US-10-007-700-347

Query Match 34.6%; Score 674.2; DB 13; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY	103	CGGTTAGGAGCGGAAATTCAAATCGAAATATTCACCCAGCTCCCGATGGGAAGTAC	162
Db	230	GGCAAGGATTCGGAACCTTCAGATACGAATATTCGCGCTCATTTACAGTGGGAGTGC	289
QY	163	TGCACAGCGCTGCGCTCAGTATGTTACAGTACAGAACTGTGAGCAAGTGAACACCGAGA	222
Db	290	TGATAGTCTTACGTCAGTATGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT	349
QY	223	GTGACAGCGAGTGTGAATGTCACTATTCACCCGAGAGAGAGAGAGAGAGAGAGAGAG	282
Db	350	CGAAACTGCAGTGTGAATGTAACTATTCAGTAAAGAGAGAGAGAGAGAGAGAGAGAG	409
QY	283	TGAGAGTGTATGCGACCGAGTGTGGAGAACATCCCTGAAAGTCTCTTACATCCCGATG	342
Db	410	ACAACTGAATGGAATTCAGTTAGAGAAATTCACCTTGAAGTAGCTATATCCCTGATG	469
QY	343	AGCAGATAGCACAGGACCTGAGAAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGC	402
Db	470	AAACGGCGCCAGCAAAACCCCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
QY	403	CCGCGCAGGGCTCACTGTGGAGAGGGGGCCCCAGCCAGAGAGAGAGAGAGAGAGAGAG	462
Db	530	GGGGCTCCTCAGGCGAGGGGTCTCCAGGATCCGTATCCAGAGAGAGAGAGAGAGAGAG	589
QY	463	CCCTTCGGCTCTGCTGCTCCACAGTATGTGGGTGCCATTTATGCAAGAGAGGGGCGCA	522
Db	590	CTCTGGCTCTGCTGCTCCACCAATTTGTGGAGCCATCATAGGAAAGAGAGAGAGAGAG	649
QY	523	CCATCCGCAACATCACAAACAGACAGCCAGTCCAAAGATAGAGTGCATAGAGAGAGAGAG	582
Db	650	CCATTGGAACATCACAAACAGACAGCCAGTCTAAATTCGATGTCACCGTAAGAAATAG	709
QY	583	CAGTGCAGTGAAGAGCCATCAGTGTGACCTCCACCCCTGAGGGCTGTCTCTCGCTT	642
Db	710	CGGGGCTGTGAGAGTGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	769
QY	643	GTAAGATGATCTGGAGATTCATTAAGAGAGCTAAGACACCAAAACGGCTGACAGG	702
Db	770	GTAAGTCTATTCTGGAGATTATGCAATGAAGAGCTCAAGATATAAAATTCACAGAGAGA	829

RESULT 11

US-10-117-982-347
; Sequence 347, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.

QY	703	TTCCCTGAAGATCTCTGGCCCATTAATAAATTGTTAGGGCTCTCATTTGGCAAGAGGAC	762
Db	830	TCCTCTTGAAGATTTTAGCTCATAATAAATTGTTGGACGCTTATTTGGTAAAGAGAA	889
QY	763	GGAACTGTGAAGAGTAGAGATACCGAGACAAATAACCAATCTCTCTCTCTCTCTCTCT	822
Db	890	GAAATCTTAAATAATTTGAGCAAGACACAGACATAAATCAGATATCTCCATTTGAGG	949
QY	823	ACCTTACCTTTTACAACCCCTGAGAGACCATCACTGTGAAGGGGGCCATCAGAAATGTT	882
Db	950	AATTGAGCTGTATATCCAGAACGACATTTACAGTTAAAGGCAATGTTGACATGTC	1009
QY	883	CGAGGCCGAGAGAGAAATTAAGAAAGTTCGGAGGCTTATGAGATGATGTTGCTG	942
Db	1010	CCAAAGCTGAGGAGAGATCATGAAGAAATCAGGGAGTCTTATGAAAATGATATTGCTT	1069
QY	943	CCATGA-----GCTCTCACTGATCCCTGGCTGACCTGGCTGCTGAGGCTTTTCC	936
Db	1070	CTATGATCTTCAAGCACATTTAATTCCTGGATTAATCTGAACGCTTGGGCTCTGTTCC	1129
QY	997	CAGCTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATA	1056
Db	1130	CACCCACTTCAGGAGTCCACCTCCAGGCGCCCTTCAGCCATGACTCTCCCT	1189
QY	1057	GCTCCTTTATGAGGCTCCCGAGCAGAGATGTTGAGTGTGTTATTCCTCCGCGCAGGAG	1116
Db	1190	ACCCGAGTTTGAGCAATCAGAAACGAGACTGTTCTATCTGTTATCCAGCTCTATCAG	1249
QY	1117	TGGCGCCCATCATCGGCAAGAGGGGCGAGCACATCAACAGCTCTCCCGTTTGCAGCG	1176
Db	1250	TCGCTGCCATCATCGCAAGAGGCGGCGAGACATCAAGCAGCTTCTCGCTTTGCTGGAG	1309
QY	1177	CTTCATCAAGATTCACCAACCGAAACCTGACTCCAAAGTTCGTATGTTTATCATCA	1236
Db	1310	CTTCAATTAAATTTGCTCCAGCGGAAGCACAGATGCTAAAGTGAGGATGGTGATTATCA	1369
QY	1237	CTGGACCGCAGAGGCCCAATTCAGGCTCAGGAGAGATCTATGCAAACTCAAGAGG	1296
Db	1370	CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGAGAGAAATTTATGGAATAATTAAGAAG	1429
QY	1297	AGAACTTCTTGTCTCCAAAGAGGAGTGAAGCTGAGAGCCACATACATGTCGCCAGCAT	1356
Db	1430	AAACTTTGTTAGTCTTAAAGAGAGTGAAGTGAAGTCTATATCAAGTGCCTCTCT	1489
QY	1357	CAGAGCTGCGCGGCTTATGGCAAGTGGAAAAACCGTGAACGAGTTGCAAGATTTGA	1416
Db	1490	TTGCTCTGCGCAGAGTTATTTGAAAAAGGAGGCAAAACCGTGAATGAATTCAGAAATTTGT	1549
QY	1417	CGGAGCTCAGTGTGTAGTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1476
Db	1550	CAAGTGCAGAGTGTGTGCTCTGCTGAGCAGACCTGATGAGAAATGCAAGTGGTTG	1609
QY	1477	TGAAATCATCGGACATTTCTATGCGAGTTCAGATGGCTCAACGGAAGATCCGAGACATCC	1536
Db	1610	TCAAAATACTGTCATCTTATGCTTGCAGGTTGCCAGAGTGGCCAGAGAAAAATTCAGAAATTC	1669
QY	1537	TGCCCCAGGTTAAGCAGAGCATCAGAG	1565
Db	1670	TGACTCAGGTAAGCAGCACCACCAACACAG	1698

Db 1610 TCAATAAAGTGGTCACTTCTATGCTTGCAGGTTGCCAGAGAAAATTCAGGAATTC 1669
QY 1537 TGGCCAGGTTAAGCAGCAGATCAGAG 1565
Db 1670 TGACTCAGGTAAGCAGCAGCAACACAG 1698

RESULT 9
US-09-897-778-347
; Sequence 347, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-778-347

Query Match 34.6%; Score 674.2; DB 9; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGAGCCGGAATTCNAATCCGAATATTCACCCAGCTCCGATGGGAAGTAC 162
Db 230 GCGAAAGATTGCGAAATTCAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGC 289
QY 163 TGGACAGCTGCTGCTCAGTATGTACAGTAGAAGTGTGAGCAAGTGAACACCGAGA 222
Db 290 TGGATAGTTACTAGTCCAGTATGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
QY 223 GTGACAGCGCAGTGGTGAATGCACTATTCCAAACCGGAGCAGACCGAGCCATCA 282
Db 350 CGGAAACTGCACTTGTAAATGTAACCTATTCCAGTAAAGCAACCAAGCTAGACAAGCACTAG 409
QY 283 TGAAGCTGAATGGCCACAGTTGGAGAACCATGCTTGAAGTCTCTACATCCCGATG 342
Db 410 ACAAACTGAATGGATTTTCAAGTAGAGAAATTCACCTTGAAGTAGCTATATCCCTGATG 469
QY 343 AGCAGATAGCAGAGGACCTGAGAATGGCGCGGAGGGGCTTTGCTCTCGGGGTGACG 402
Db 470 AAACGCCGCCAGCAAAACCCCTTGCAGACGCCCGAGTCCGCGGGGCTTGGGCGAGA 529
QY 403 CCGCGCAGGGCTCAGCTGTGGAGCGGGGGCCCCAGCAGCAGCAGCAAGTGGACATCC 462
Db 530 GGGGCTCCTCAAGCAGGGGCTTCCAGGATCCGATATCCAAAGCAGAAACCATGTGATTCG 589
QY 463 CCCTTCGGCTCTGTTGCCACCCAGTAGTGGGTGCCATTTATGGCAAGGAGGGGCCA 522
Db 590 CTCTGCGCTGCTGGTTCCACCCAAATTTTGGAGCCATCATAGAAAGAGGTGCCA 649
QY 523 CCATCCGCAACATCAAAAACAGACCCAGTCCAAAGATAGACGTGCTATAGAGGAGAACG 582
Db 650 CAAATTCGGACATCACCMAACAGACCCAGTCAAAATCGATGTCCACCGTAAAGAAATG 709
QY 583 CAGGTCCAGTGAAGAACCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTT 642

Db 710 CGGGGCTGCTGAGAGTCGATTACTATCTCTCTACTCTCTGAAAGCACCCTCTCGGGCTT 769
QY 643 GTAAGATGATCTTGGAGATTTATGATATAAGAGGCTAAGGACACCAAAACGCTGACGAGG 702
Db 770 GTAAGTCTATTTCTGGAGATTTATGATATAAGGAAGCTCAAGATATAAAATTTCAAGAAGA 829
QY 703 TTCCCTGGAAGATCTCGCCCATATAACTTTTGGGCGCTCTCATTTGGCAAGGAAGAC 762
Db 830 TCCCTTTGAAGATTTTAGCTCATATAACTTTTGGACGCTTTATTTGGTAAGAGGNA 889
QY 763 GGAACCTGGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTGCAAG 822
Db 890 GAAATCTTAAATAAATTTGAGCAAGACACAGACATAAAATCACGATATCTCAITTCAGG 949
QY 823 ACCTTACCTTTACAACCTTGAGAGGACCATCAGTGAAGGGGGCCATCAGAGATTCCTT 882
Db 950 AATTGACGCTGTATAATCCAGAACGCTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009
QY 883 GCAGGGCGGAGCAGAAATATGAAGAAAGTTTCGGAGGCTATGAGAAATGATGCGGTG 942
Db 1010 CCAAGCTGAGGAGAGATCATGAAGAAATCAGGAGTCTTATGAAAATGATATGCTT 1069
QY 943 CCATGA-----GCTCTCACCTGATCCCTGGCCTGAACCTGGCTGCTAGCTCTTTTCC 996
Db 1070 CTATGAATCTTCAAGCACATTTAATCTCTGATTAAATCTGAACGCTTTGGTCTGTTC 1129
QY 997 CAGCTTCAATCCAGCGCAGTCCCGCGCTCCAGCAGAGGTTACTGGGGCTGCTCCCTATA 1056
Db 1130 CACCACTTCAGGGATGCCACTCCCACTCAGGGGCCCCCTTCAGCCATGACTCTCCCT 1189
QY 1057 GCTCTTTATGACGCTCCCGAGCAGAGATGTTGTCAGTGTATATCCCGCCAGGAGG 1116
Db 1190 ACCCGAGTTTGAGCAATCAGAAACGAGACTGTTTCATCTGTTATCCAGCTCTATCAG 1249
QY 1117 TGGCGCCATCATCGCAAGAGGGGCGAGCATCAAAACAGCTCTCCCGTTTGGCAGG 1176
Db 1250 TCGGTGCATCATCGCAAGCAGGGCCAGCATCAAGCAGCTTTCTCGCTTTGCTGGAG 1309
QY 1177 CTTCAATTAAGATTGCTCCAGCGAAGCAGATGCTAAAGTGAGGATGGTGAATATCA 1369
Db 1237 CTGACCGCCAGAGCCCAATTCAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGG 1296
QY 1370 CTGACCCACAGAGCTCAGTTCAAGCTCAGGGAAGATTTATGGAAAAATTAAGAAG 1429
QY 1297 AGAACTTCTTTGCTCCCAAGGAGGAGTGAAGCTCGAGACCCACATACGTGTGCCAGAT 1356
Db 1430 AAACTTTTGTAGTCTTAAAGAGAGGTGAATCTTGAAGCTCATATCAGAGTGCATCCT 1489
QY 1357 CAGCAGTGGCGCGGTCAATTTGGCAAGAGTGGAAAAACGTTGAACGAGTTGCAGATTTGA 1416
Db 1490 TTGCTGCTGGCAGAGTTTATGGAAAAAGGAGGCAAAACGTTGAATGAATTCAGAAATTT 1549
QY 1417 CGGCAAGCTGAGTGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACCCAGGTCATCG 1476
Db 1550 CAAGTGCAGAAATTTGTGCTCTGTCGACCAACACCTGATGAGATGACCAAGTGGTTG 1609
QY 1477 TGAATAATCATCGCAATTTCTATGCGAGTTCAGATGGCTCAACCGAAGATCCGAGACATCC 1536
Db 1610 TCAATAAATAGTGTCTCATCTTATGCTTGCAGGTTGCCAGAGGAAAAATTCAGGAAATTC 1669
QY 1537 TGGCCAGGTTAAGCAGCAGATCAGAG 1565
Db 1670 TGACTCAGGTAAGCAGCAGCAACACAG 1698

RESULT 10
US-10-007-700-347
; Sequence 347, Application US/10007700
; Publication No. US2003064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong

Db 1310 CTTCAATTAAAGATTGCTCCAGCGGAGACACAGATGCTAAAGTGAGGATGGTGAATTACA 1369
Qy 1237 CTGAGACCGCCAGAGGCCCAATTAAGGCTCAGGGAAGAATCTATGGCAAACTCAAGGAGG 1296
Db 1370 CTGGACCAACAGAGGCTCAGTTCAGGCTCAGGGAAGAATTTATGGAAAATTAAGAAG 1429
Qy 1297 AGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTTGGAGACCCACATAGCTGTGCCAGAT 1356
Db 1430 AAAACTTTGTAGTCTTAAGAAGAGGTGAACCTTGAAGCTCATATCAGAGTGCCATCCT 1489
Qy 1357 CAGCAGCTGGCCGGGTCAATTTGGCAAAAGTGGAAAACGGTGAACAGTTCAGAAATTTGA 1416
Db 1490 TTGCTGCTGCGAGAGTTATTGGAAAAGGAGCAAAACGGTCAATGAATTCAGAAATTTGT 1549
Qy 1417 CGGCACTGAGGTGGTAGTACCAAGAGACACAGACCCCTGATGAGAAACGACAGGTCAATCG 1476
Db 1550 CAAATGCAGAAAGTTGTTGCTCCCTGTCAGCAGACACCTGATGAGAAATGACCAAGTGGTTG 1609
Qy 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCGAGACATCC 1536
Db 1610 TCAATAATCACTGTCATCTTATGCTTGCAGGTGGCCAGAGAAAATTCAGGAAAATTC 1669
Qy 1537 TGGCCCAAGGTTAAGCAGCAGCATCAGAAAG 1565
Db 1670 TGACTCAGGTAAGCAGCAGCAACCAACAG 1698

RESULT 8

US-09-850-716A-347
; Sequence 347, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-347

Query Match 34.6%; Score 674.2; DB 9; Length 1740;
Best Local Similarity 66.7%; Pred. No. 6.6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
Qy 103 GGGTTAGGAGCGGAAATTCAAATCCGAATATCCACCCAGCTCCGATGGGAAGTAC 162
Db 230 GGCAAAAGGATTCGGAACATTCAGATACGAATATCCCGCTCATTTACAGTGGGAGTGC 289
Qy 163 TGACAGCCTGCTGGCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGA 222
Db 290 TGGATAGTTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
Qy 223 GTGAGACGGCAGTGGTGAATGTCACTATTCCAAACCGGAGCAGACACGAGCCATCA 282
Db 350 CGGAAACTGCAGTTGTAATGTAACCTATTCAGTAGAGACCAAGCTAGACAGCACTAG 409
Qy 283 TGAAGCTGAATGCCACAGATTGGAGAACATGCTCCCTGAAGTCTCTTACATCCCGATG 342
Db 410 ACAAACTGAATGGAATTCAGTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG 469
Qy 343 AGCAGATAGCAGAGGACCTGAGAAATGGGCGCAGGGGCTTTGGCTCTCGGGGTACG 402
Db 470 AAACGCCGCCAGCAAAACCCCTTCGACAGACGCCCGAGGTGCGCGGGGGCTTGGGCAGA 529

Qy 403 CCCGCCAGGGGCTACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCC 462
Db 530 GGGGCTCCTCAAGGCAAGGGGTCTCCAGGATCCGATATCCAAGCAGAAACCATGTGATTGC 589
Qy 463 CCCTTGGGCTCCTGGTGGCCACCCAGTATGTGGTGGCATTTATGGCAAGAGGGGGCCA 522
Db 590 CTCTGCGCCCTGCTGGTGTCCCAACCCCAATTTGTGAGCCATCATAGGAAAAGAGGTGCCA 649
Qy 523 CCATCCGCAACATCACAAAAACAGACCCAGTCCCAAGATAGACGTGCAATAGGAAGAGAACG 582
Db 650 CCATTGGGAACATCACCAACAGACCCAGTCTAAATCGATGTCCACCGTAAAGAAATG 709
Qy 583 CAGGTGACGTGAAAAGCCATCAGTGTGCCTCCACCCCTGAGGGGTGTCTCTCCGCTT 642
Db 710 CGGGGGCTGCTGAGAAGTCGATTAATCTCTCTACTCTCTGAAGGCACCTCTCTGGGCTT 769
Qy 643 GTAAGATGATCTTGGAGATTATGATAAAGAGCTAAGGACCAAAACGGCTGACAGG 702
Db 770 GTAAGTCTATCTGGAGATTATGCATAGGAAGCTCAAGATATAAAATTCACAGAGAGA 829
Qy 703 TTCCCTGAAGATCCTGGCCCATATAAATCTTTGTAGGGCTCTCATTTGGCAAGAAAGGAC 762
Db 830 TCCCTTGAAGATTTTAGCTCATATAACTTTGTGACGTCTTATTGGTAAAGAGGAA 889
Qy 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCATCTCTCTCTGCAAG 882
Db 890 GAAATCTTAAAAAAATTTGAGCAAGACACACACTAAATACGATATCTCCATTGCGAG 949
Qy 823 ACCTTACCCCTTTACAAACCTTGAGAGACCATCACTGTGAAGGGGCCATCGAGAATGTT 882
Db 950 AATTGAGCTGTATATCCAGAAGCAGCATTAACAGTTAAAGCAATTTGAGACATGTG 1009
Qy 883 GCAGGCGGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCTATGAGAATGATGTGGCTG 942
Db 1010 CCAAAGCTGAGGAGAGATCATGAAGAAAATCAGGGAGTCTTATGAAAATGATATTGCTT 1069
Qy 943 CCAATGA-----GCTCTCACTGATCCCTGGCTGAACTGGCTGTGTAGTCTTTTCC 996
Db 1070 CTATGAATCTTCAAGCACATTTAAATCTTGGATTTAAATCTGAACGCTTTGGGTCTGTCC 1129
Qy 997 CAGCTTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGGTTACTGGGGCTGCTCCCTATA 1056
Db 1130 CACCCACTTCAGGATGCCACCTCCACCTCAGGGCCCTTCAGCCATGACTTCTCCCT 1189
Qy 1057 GTCCTTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTTATCCCGCCCGCAGCAG 1116
Db 1190 ACCCGCAGTTTGAGCAATCAGAAAACGAGACTGTTTCATCTGTTTATCCAGCTCTATCAG 1249
Qy 1117 TGGGGCCATCATCGGCAAGAGGGGCGACGACATCAACAGCTCTCCCGTTTGCAGCG 1176
Db 1250 TCGTGCCCATCATCGGCAAGCGGGCCAGCATCAAGCAGCTTTCTCGCTTTGTGCGAG 1309
Qy 1177 CTTCCATCAAGATTCACCAACCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCATCA 1236
Db 1310 CTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTTAAAGTGAAGTGGTGAATTATCA 1369
Qy 1237 CTGAGACCGCCAGAGGCCCAATTAAGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAG 1296
Db 1370 CTGGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAAAATTAAGAAG 1429
Qy 1297 AGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGAT 1356
Db 1430 AAAACTTTGTAGTCTTAAAGAGAGGTGAAACTTTGAAGCTCATATCAGAGTGCCATCCT 1489
Qy 1357 CAGCAGCTGGCCGGGTCAATTTGGCAAAAGTGGAAAACCGTGAACGAGTTGCAAGATTTGA 1416
Db 1490 TTGCTGCTGGCAGAGTTATTGAAAAGAGGCAAAACGGTGAATGAATTCAGAAATTTGT 1549
Qy 1417 CGGCACTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACCAAGGTCAATCG 1476
Db 1550 CAAAGTGCAGAGTGTGTTGTCTCTGTCGACACACCTGATGAGAAATGACCAAGTGGTTG 1609
Qy 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536

541 AAAACGGTGAACAGATTGCGAATTTGACCGGAGCTGAGTGGTAGTACCAAGAGACAG 600
1449 ACCCTCATGAGAACAGCAGGTCATCGTGAATAATCATCGACACATTTCTATGCCAGTCAG 1508
601 ACCCTCATGAGAACAGCAGGTCATCGTGAATAATCATCGACATTTCTATGCCAGTCAG 660
1509 ATGGCTCAACCGGAAGATCCGAGACATCCTGGCCAGGTTTAAAGCAGCGATCAGAAGGGA 1568
661 ATGGCTCAACCGGAAGATCCGAGACATCCTGGCCAGGTTTAAAGCAGCGATCAGAAGGGA 720
1569 CAGATTAACAGGCCCGGACCGAGAGTGAACAGCCCTCCCTGTCCTTTGAGTCC 1628
721 CAGATTAACAGGCCCGGACCGAGAGTGAACAGCCCTCCCTGTCCTTTGAGTCC 780
1629 AGGACAAACCGGCAGAAATCGAGAGTGTCTCTCCCGGAGCGCTTGAGAAATCAGTGG 1688
781 AGGACAAACCGGCAGAAATCGAGAGTGTCTCTCCCGGAGCGCTTGAGAAATCAGTGG 840
1689 GAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTTGGCCATTTGAGAAAGATG 1748
841 GAATCCGGGACACNTGGGCCGGGCTGTAGATCAGGTTTGGCCATTTGAGAAAGATG 900
1749 TTCAGTGAAGAACCTGATCTNTCAGCCCAACACCCCAATTTGGCCCAACACATGT 1808
901 TTCAGTGAAGAACCTGATCTCTCAGCCCAACACCCCAATTTGGCCCAACACATGT 960
1809 NTGCCCTCGGGGTGTGAGAAATTTAGCGAAGGCACTTTTAAACGTGGATTGTTTAAA 1868
961 CTGCCCTCGGGGTGTGAGAAATTTAGCGAAGGCACTTTTAAACGTGGATTGTTTAAA 1020
1869 GAAGCTCTCCAGGCCCGCCACCAAGAGGGTGGATTCACCTCAGTGGGAAAGAAATATAAT 1928
1021 GAAGCTCTCCAGGCCCGCCACCAAGAGGGTGGATTCACCTCAGTGGGAAAGAAATATAAT 1080
1929 TTCCTTCAGGTTTAAAA 1946
1081 TTCCTTCAGGTTTAAAA 1098

RESULT 7
US-09-735-705-347
; Sequence 347, Application US/09735705
; Patent No. US2002052329A1

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Ligu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-705-347

Query Match 34.6%; Score 674.2; DB 9; Length 1740;
Best Local Similarity 56.7%; Pred. No. 6 6e-200;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;

QY 103 GGGTTAGGAGCGGAAAAATTCAAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTAC 162
DB 230 GGCMAAGGATTCGGAATCTTCAGATACGAATATCCCGCTCATTTACAGTGGGAGTGC 289
QY 163 TGGACAGCCTCTGGCTCAGTATGGTACAGTAGAAGACTGTGAGCAGTGAACCCGAGA 222
DB 290 TGGATAGTTTACTAGTCCAGTATGGAGTGGGAGAGCTGTGAGCAGTGAACCTGACT 349
QY 223 GTGAGACGGAGTGGTGAATGTACCTTACCTTATCCACCGGGAGCAGACCAAGCAATCA 282
DB 350 CGGAAACTGCAAGTTGTAATGTAACCTTATCCAGTAAGGACCAAGCTAGAACGACCTAG 409
QY 283 TGAAGCTGAATGGCCACACAGTTGAGAACCATGCCCTGAAGGCTCTCTACATCCCGGATG 342
DB 410 ACAACTGAATGGATTTTCAGTTAGAGAAATTCACCTTGAAGTAGCCTATATCCCTGATG 469
QY 343 AGCAGATAGCACAGGAGCCTGAGAAATGGGCCCGAGGGGGCTTTGGCTCTCGGGGTACG 402
DB 470 AAACGGCGCCCGACAAAACCCCTTGACAGCAGCCCGAGGTCGCGGGGGCTTGGGCAGA 529
QY 403 CCGCCAGGGCTCAGCTGTGSCAGCGGGGGCCCGAGCCCAAGCAGCAGCAAGTGGACATCC 462
DB 530 GGGGCTCTCTCAAGGACGGGCTCTCCAGGATCCGATCCGATCCAGGAGAAACCATGTGATTGC 589
QY 463 CCCTTCGGCTCTCTGGTCCCGACCCAGTATGTGGGTGCCATTTATGGCAAGGAGGGGCCA 522
DB 590 CTCTGGGCTCTCTGGTCCCGACCCAAATTTGTGGAGCCATCATAGGAAAGAGGTGCCA 649
QY 523 CCATCCGCAACATCACAAACAGACCCAGTCCAGATAGACGTGATGAGAGGAGAGAGCG 582
DB 650 CCATTCGGAACATCACCAACAGACCCAGTCTAAATCGATGTCCACCGTAAAGAAAATG 709
QY 583 CAGGTGAGCTGAAAGAACCCATCAGTGTGCACTCCACCCCTGAGGGCTCTCTCCGCTT 642
DB 710 CGGGGCTCTCTGAGAGTTCGATTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
QY 643 GTAAGATCATCTCGAGATTTATGCATAAAGAGGCTTAAGGACACCAAAACGGCTGACGAG 702
DB 770 GTAAGTCTATTCTGGAGATTATGCATAAGAGAGCTCAGATATAAATTCACAGAGAGA 829
QY 703 TTCCCTCGAAGATCCTGGCCCATATAAATTTGTAGGGGCTCTCATTTGGCAAGGAGAGAC 762
DB 830 TCCCTCTGAGATTTTAGCTCATATAAATTTGTGGAGCTCTTATTTGTAAGAAGAGAA 889
QY 763 GGAACCTGAAGAGGTAGACGAGATACCGAGACAAATACCATCTCTCTCTCTCTCTCT 822
DB 890 GAAATCTTAAAAAATTTGAGCAAGACACAGACATAAATACGATATCTCCATTTGCAGG 949
QY 823 ACCTTACCTTTTACAACTGAGAGGACCATCACTGTGAAAGGGGCGCATTCGAGAAATTT 882
DB 950 AATTGACGCTGTATAATCCAGAACGCACTTATACAGTTAAAGGCAATGTTGAGACATGT 1009
QY 883 GCAGGGCGGAGCAGGAAATTAATGAAGAAAGTTTCGGAGGCGCTATGAGATGATGGCTG 942
DB 1010 CCAGAGCTGAGGAGGAGATCATGAAGAAATCAGGAGTCTTATGAAATGATTAATTTCT 1069
QY 943 CCATGA-----GCTCTACCTGATCCCTGGCTGACCTGCTGCTGCTGCTGCTTTTCC 996
DB 1070 CTATGAATCTTCAAGCACATTTAATTCCTGGATTAATTAATGAACGCTTTGGGCTCTG 1129
QY 997 CAGCTTCATCCAGCGCAGTCCCGCGCTCCCGAGCAGCGCTTACTTGGGGCTGCTCCCTATA 1056
DB 1130 CACCACTTCAGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGACTCTCCCT 1189
QY 1057 GCTCCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGTGTGTTTATCCCGCCCGAGGAG 1116
DB 1190 ACCCGCAGTTTCAGCAATCAGAAAACGAGACTGTTTCATCTGTTTATCCAGCTCTATCAG 1249
QY 1117 TGGGGCCCATCATCGCAGAGAGGGGCGACATCAACAGCTCTCCCGGCTTGGCCAGCG 1176
DB 1250 TCGGTGCCATCATCGGCAAGAGGGGCGACATCAAGAGCTTTTCTCGTTTCTGCTGGAG 1309
QY 1177 CTTCCATCAAGATTGACCAACCGCAACACCTGACTCCAAAGTTCGTATGTTTATCATCA 1236

; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-764-864-327									
Query Match 55.8%; Score 1085.4; DB 9; Length 1182;									
Best Local Similarity 98.9%; Pred. No. 0;									
Matches 1102; Conservative 0; Mismatches 6; Indels 6; Gaps 1;									
QY	839	CCCTGAGAGACCATCTCTGTGAAGGGGGCCATCGAAGATTGTTCAGGGCCGAGCAGGA	898						
Db	18	CGCTGAGAGACCATCATCTGTGAAGGGGGCCATCGAAGATTGTTCAGGGCCGAGCAGGA	77						
QY	899	AATAATGAAGAAATTCCGGAGGCTATGAGATGATGCTGCCATGAGC-----TC	952						
Db	78	AATATGAGAGAGTTTCGGAGGGCTATGAGATGATGCTGCCATGAGCCTTCAGTC	137						
QY	953	TCACCTGATCCCTGGCTGAACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGC	1012						
Db	138	TCACCTGATCCCTGGCTGAACCTGGCTGCTGTAGGTCTTTTCCAGCTTCATCCAGCGC	197						
QY	1013	AGTCCCGCGCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGCTTCCTTTATGAGGC	1072						
Db	198	AGTCCCGCGCTCCAGCAGCGTTACTGGGGTGTCTCCCTATAGCTTCCTTTATGAGGC	257						
QY	1073	TCCCGAGCAGGATGGTGAGTGTGTTTATCCCGCCAGGCGAGTGGGGCCATCATCGG	1132						
Db	258	TCCCGAGCAGGATGGTGAGTGTGTTTATCCCGCCAGGCGAGTGGGGCCATCATCGG	317						
QY	1133	CAAGAGGGGACGACATCAAAACAGCTCTCCCGGTTTGGCAGGCGCTCCATCAAGATTGC	1192						
Db	318	CAAGAGGGGACGACATCAAAACAGCTCTCCCGGTTTGGCAGGCGCTCCATCAAGATTGC	377						
QY	1193	ACCACCGGAAACCTGACTCCAAAGTTCGTATGGTTATCATCTGAGCGCCGACAGGC	1252						
Db	378	ACCACCGGAAACCTGACTCCAAAGTTCGTATGGTTATCATCTGAGCGCCGACAGGC	437						
QY	1253	CCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGGAACTTCTTTGGTCC	1312						
Db	438	CCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGGAACTTCTTTGGTCC	497						
QY	1313	CAAGGGAAGTGAAGCTGGAGCCACATACGTGTGCCAGCATCAGCAGCTGGCCGGGT	1372						
Db	498	CAAGGGAAGTGAAGCTGGAGCCACATACGTGTGCCAGCATCAGCAGCTGGCCGGGT	557						
QY	1373	CATTGCAAGGTGGAAACGTTGACGATTTGCAAAATTCACGGCAGCTGAGTGGT	1432						
Db	558	CATTGCAAGGTGGAAACGTTGACGATTTGCAAAATTCACGGCAGCTGAGTGGT	617						
QY	1433	AGTACCAAGAGACGAGACCCCTGATGAGAACGACAGGTTCATCGTGAATAATCATCGACA	1492						
Db	618	AGTACCAAGAGACGAGACCCCTGATGAGAACGACAGGTTCATCGTGAATAATCATCGACA	677						
QY	1493	TTTCTATGCGATCAGATGGCTCAACGGAGATCCGAGACATCCTGGCCGAGTTAAGCA	1552						
Db	678	TTTCTATGCGATCAGATGGCTCAACGGAGATCCGAGACATCCTGGCCGAGTTAAGCA	737						
QY	1553	GCAGCATCAGAGGGACAGATTAAACAGGCCAGGCAACGAGGAATGACAGCCCTCC	1612						
Db	738	GCAGCATCAGAGGGACAGATTAAACAGGCCAGGCAACGAGGAATGACAGCCCTCC	797						
QY	1613	CTGTCCCTTNGATCCAGGACAAACAGGGCAGAAATCGAGAGTGTGCTCTCCCGGCGAG	1672						
Db	798	CTGTCCCTTNGATCCAGGACAAACAGGGCAGAAATCGAGAGTGTGCTCTCCCGGCGAG	857						
QY	1673	GCCTGAGATAGTGGGAATCCGGGACACNTGGCCGGGCTGTPAGATCAGGTTTGCCAC	1732						
Db	858	GCCTGAGATAGTGGGAATCCGGGACACNTGGCCGGGCTGTPAGATCAGGTTTGCCAC	917						
QY	1733	TTGATTGAGAAAGATGTTCCAGTGGAGAACCTTGATCTTTCAGCCCAACACCCACCA	1792						
Db	918	TTGATTGAGAAAGATGTTCCAGTGGAGAACCTTGATCTTTCAGCCCAACACCCACCA	977						
QY	1793	ATTGGCCCAACACTGTNTGCCCTCGGGGTGTGAGAAATTNTAGCGCAAGGCATTTTAA	1852						
Db	978	ATTGGCCCAACACTGTNTGCCCTCGGGGTGTGAGAAATTNTAGCGCAAGGCATTTTAA	1037						
QY	1853	ACGTGATTTGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTG	1912						
Db	1038	ACGTGATTTGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGGTGGATCACACCTCAGTG	1097						
QY	1913	GGAGAAAAATAAAATTTCTTCCAGGTTTAAAA 1946							
Db	1098	GGAGAAAAATAAAATTTCTTCCAGGTTTAAAA 1131							
RESULT 6									
US-09-764-864-745									
; Sequence 745, Application US/09764864									
; Patent No. US20020132753A1									
; GENERAL INFORMATION:									
; APPLICANT: Rosen et al.									
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies									
; FILE REFERENCE: PT223									
; CURRENT APPLICATION NUMBER: US/09/764,864									
; CURRENT FILING DATE: 2001-01-17									
; Prior application data removed - consult PALM or file wrapper									
; NUMBER OF SEQ ID NOS: 1792									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 745									
; LENGTH: 1129									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-09-764-864-745									
Query Match 55.0%; Score 1069.8; DB 9; Length 1129;									
Best Local Similarity 98.7%; Pred. No. 0;									
Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;									
QY	855	ACTGTGAAGGGGGCCATCGAGAAATTGTCGAGGCGGAGGAGGAATAATGAGAAAGTT	914						
Db	1	ACTGTGAAGGGGGCCATCGAGAAATTGTCGAGGCGGAGGAGGAATAATGAGAAAGTT	60						
QY	915	CGGGAGGCTTATGAGAAATGATGTGGCTGCCATGAGC-----TCTCACCTGATCCCTGGC	968						
Db	61	CGGGAGGCTTATGAGAAATGATGTGGCTGCCATGAGCTGCAGTCTCCTGATCCCTGGC	120						
QY	969	CTGAACCTGGCTGTGTAGGTCTTTTCCAGCTTCAATCCAGCGCAGTCCCGCGCTCCC	1028						
Db	121	CTGAACCTGGCTGTGTAGGTCTTTTCCAGCTTCAATCCAGCGCAGTCCCGCGCTCCC	180						
QY	1029	AGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGGCTCCCGAGCAGGATG	1088						
Db	181	AGCAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGGCTCCCGAGCAGGATG	240						
QY	1089	GTGCAAGTGTATATCCCGCCCGAGGCGAGTGGCGCCATCATCGCAAGAGGGCGAGCAC	1148						
Db	241	GTGCAAGTGTATATCCCGCCCGAGGCGAGTGGCGCCATCATCGCAAGAGGGCGAGCAC	300						
QY	1149	ATCAAAACAGCTCTCCGGTTTTCGAGCGCTCCATCAAGATTGACACCCGAAACACCT	1208						
Db	301	ATCAAAACAGCTCTCCGGTTTTCGAGCGCTCCATCAAGATTGACACCCGAAACACCT	360						
QY	1209	GACTCCAAAGTTCGTATGGTATCATCTGACCGCCGAGGCGCCAAATTCAGGCTCAG	1268						
Db	361	GACTCCAAAGTTCGTATGGTATCATCTGACCGCCGAGGCGCCAAATTCAGGCTCAG	420						
QY	1269	GGAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGAGGAAAGTGAAG	1328						
Db	421	GGAGAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGAGGAAAGTGAAG	480						
QY	1329	CTGGAGACCCACATACGTGTGCGCAGCATCAGAGTGCCTGGGTCAATGGCAAAAGTGA	1388						
Db	481	CTGGAGACCCACATACGTGTGCGCAGCATCAGAGTGCCTGGGTCAATGGCAAAAGTGA	540						
QY	1389	AAAAACGGTGAACGAGTTCGAGAAATTTGACGCGAGCTGAGTGGTAGTACCAAGACACAG	1448						

; SOFTWARE: PERL Program									
; SEQ ID NO 2082									
; LENGTH: 1738									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; NAME/KEY: misc feature									
; OTHER INFORMATION: Incyte ID No. US20040005559A1 457674.1									
US-10-062-674-2082									
Query Match 66.4%; Score 1291.4; DB 16; Length 1738;									
Best Local Similarity 93.8%; Pred. No. 0;									
Matches 1621; Conservative 0; Mismatches 61; Indels 46; Gaps 25;									
QY	265	AGACGAGCGAGCCATCATGAAGCTGAATGCGCCACCACTGTTGGAGAACCATCCCTCGAAGG	324						
DB	1738	AGACGAGCGAGCCATCATGAAGCTGAATGCGCCACCACTGTTGGAGAACCATCCCTCGAAGG	1679						
QY	325	TCTCTACATCCCGATGAGCAGATAGCACAGGACCTTGAGAAATGGGGCGCGAGGGGCT	384						
DB	1678	TCTCTACATCCCGATGAGCAGATAGCACAGGACCTTGAGAAATGGGGCGCGAGGGGCT	1619						
QY	385	TTGGCTCTCGGGTTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGC	444						
DB	1618	TTGGCTCTCGGGTTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGC	1559						
QY	445	AGCAGCAAGTGACATCCCTCTCGGCTTCCTGGTGGCCACCCAGTATGTGGGTGCCATTA	504						
DB	1558	AGCAGCAAGTGACATCCCTCTCGGCTTCCTGGTGGCCACCCAGTATGTGGGTGCCATTA	1499						
QY	505	TTGGCAGGAGGGGGCCACCATCCGCAACATCACAAAAAGAGCCAGTCCCAAGTAGAGG	564						
DB	1498	TTGGCAGGAGGGGGCCACCATCCGCAACATCACAAAAAGAGCCAGTCCCAAGTAGAGG	1439						
QY	565	TGCATAGAGAGAGAACCGAGCTGCAGCTGAAAGAGCCATCAGTGTCACTCCACCCCTG	624						
DB	1438	TGCATAGAGAGAGAACCGAGCTGCAGCTGAAAGAGCCATCAGTGTCACTCCACCCCTG	1379						
QY	625	AGGGCTGCTCTCCG-CTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGAC	683						
DB	1378	AGGGCTGCTCTCTCCG-CTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGAGCTAAGAC	1319						
QY	694	ACCAAAACGGCTGACGAGGTTCCCTGAAAGATCTTGGGCCATTAATACTTTGTAGGGGCT	743						
DB	1318	ACCAAAACGGCTGACGAGGTTCCCTGAAAGATCTTGGGCCATTAATACTTTGTAGGGGCT	1259						
QY	744	CTC-ATTGGGGAAGAGACGGAACCTGAAGAA-GGTAGAGCAAGATACCGAGACAAAA	801						
DB	1258	CTCAATTGGCAAGGAAGACGGAACCTGAAGAAAGGTTAGACATACCGAGACAAAA	1199						
QY	802	TCACCATCTCTCTGTTGCAAGACCTTACCTTTTCAA-CCCTGAGAGGACCATCACTGTG	860						
DB	1198	TCACCATCTCTCTGTTGCAAGACCTTACCTTTTCAA-CCCTGAGAGGACCATCACTGTG	1139						
QY	861	AA-GGGGGCCATCGAGAAATTTGACAGG-CCGACAGGAAATAATGAAGAAATTCGGG	918						
DB	1138	RAAGGGGGCCATCGAGAAATTTGACAGG-CCGACAGGAAATAATGAAGAAATTCGGG	1079						
QY	919	AGGCTATGAGAAATGATGGTGTGCAATGAGC-----TCTCACTGATCCCTGGCGTGA	972						
DB	1078	AGGCTATGAGAAATGATGGTGTGCAATGAGC-----TCTCACTGATCCCTGGCGTGA	1019						
QY	973	ACCTGGCTCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGCCCTCCGAGCA	1032						
DB	1018	ACCTGGCTCTGTAGTCTTTTCCAGCTTTCATCCAGCGCAGTCCCGCGCCCTCCGAGCA	959						
QY	1033	GGGTACTGGGGTGTCTCTATAGTCTCTTTATGCAAGCTCCCGAGGAGGATGGTGC	1092						
DB	958	GGGTACTGGGGTGTCTCTATAGTCTCTTTATGCAAGCTCCCGAGGAGGATGGTGC	899						
QY	1093	AGGTGTTATCCCGCCAGGAGGAGTGGGGCGCATCATCGCAAGAGGGGCGACATCA	1152						
DB	898	AGGTGTTATCCCGCCAGGAGGAGTGGGGCGCATCATCGCAAGAGGGGCGACATCA	839						

RESULT 5
US-09-764-864-327
; Sequence 327, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 327
; LENGTH: 1182

QY	1153	AAAGCT--CTCCCGGTTTCCAGCGGCTCC-ATCAAGATTGACACCCGAAACACCTG	1209
DB	838	AAGCAGCTCTCTCCCGGTTTCCAGCGGCTCCAAATCAAGATTGACACCCGAAACACCTG	779
QY	1210	ACTCCAAAGTTGGTATGTTATCATCACTTGACCGCCAGAGG-CCCAATTCAAGGCTCAG	1268
DB	778	ACTCCAAAGTTGGTATGTTATCATCACTTGACCGCCAGAGGCCCCCAATTCAAGGCTCAG	719
QY	1269	GGAAAGAAATCTATGG-CAAACCTCAAGAGGAGAACTTCTT-TGGTCCCAAGAGGAGTGA	1326
DB	718	GGAAAGAAATCTATGGCCAAACTCAAGAGGAGAACTTCTTGGGGTCCCAAGAGGAGTGA	659
QY	1327	AGCTGGAGACCCCATATCACTGTGCCAGCATCAGCAGTGGCCGG--TCATTGGCAAGG	1384
DB	658	AGCTGGAGACCCCATATCACTGTGCCAGCATCAGCAGTGGCCGGGCTCATTTGGCCAAAGG	599
QY	1385	TGGAAAAAC-GGTCAACAGATTGACAGAAATTTGACGGCAGCTGAGTGGTAG----TACCA	1439
DB	598	TGGAAAAACGGGTGACAGATTGACAGAAATTTGACGGCAGCTGAGTGGTAGTACCAAGG	539
QY	1440	AGAGACAGACCCCTGTATGAGAACGACAGGTCTATCGTGAATAATCATCGAC-ATTTCCTA	1498
DB	538	AGACAGACCCCTGTATGAGAACGACAGGTCTATCGTGAATAATCATCGACCAATTTCTA	479
QY	1499	TGCCAGTCAGATGGCTCAACGGAA-GATCCGAGACATCTGCGCCAGGTTAAGCAGCAGC	1557
DB	478	TGCCAGTCAGATGGCTCAACGGAAAGGATCCGAGACATCTGCGCCAGGTTAAGCAGCAGC	419
QY	1558	ATCAGAGGGACAGAGTAACACAGGCCAGGACCGAGGAAGTGACAGCCCTC-CTGT	1616
DB	418	ATCAGAGGGACAGAGTAACACAGGCCAGGACCGAGGAAGTGACAGCCCTCCTGCTGT	359
QY	1617	CCCTTNGAGTCCAGGACAAACAAACGGGCGAGAAATCAGAGTGTGCTCTCCCGGAGGCT	1676
DB	358	CCCTTNGAGTCCAGGACAAACAAACGGGCGAGAAATCAGAGTGTGCTCTCCCGGAGGCT	299
QY	1677	GAGATGAGTGGGAATCCGGACAC-NTGGGCGGGCTGTAGATCAGTGTTCGCCAC----	1732
DB	298	GAGATGAGTGGGAATCCGGGACACACTGCGGCGGGCTGTAGATCAGTGTTCGCCACACT	239
QY	1733	--TTGATTTGAGAAAGATGTTCCAGTGAGAAACCTTGATCTTNTCAGCCCC-AAAA	1784
DB	238	GATTTGAGAAACGATGTTCCAGTGAGAAACCTTGATCTTNTCAGCCCCCACTACAC	179
QY	1785	CCCAACCAATTGG-CCCAACACTGTNTGCCCTCGGGGTGT-CAGAAATTTAGCGCA	1840
DB	178	CACACCAATTTGGCGCCCAACAACTGTCTGCGCTCGGGGTGTCCAGAAATCTAGCGCA	119
QY	1841	AG-GCACTTTTAAAGT-GGATTTGTTAAAGAGCTCTCCAGGCCCCCAAGAGGCTG	1898
DB	118	AGAGCACTTTTAAAGTGGGATTTGTTTAAAGAGCTCTCCAGGCCCCCAAGAGGCTG	59
QY	1899	ATCAGACCTCAGTGGGAGAGAAATAATAATTTCTTTCAGTTTAAAA 1946	
DB	58	ATCAGACCTCAGTGGGAGAGAAATAATAATTTCTTTCAGTTTAAAA 11	

; PRIOR APPLICATION NUMBER: US/09/061,709									
; PRIOR FILING DATE: 1998-04-17									
; NUMBER OF SEQ ID NOS: 8									
; SEQ ID NO 5									
; LENGTH: 1708									
; TYPE: DNA									
; ORGANISM: Homo sapiens									
; FEATURE:									
US-09-899-651-5									
Query Match 85.8%; Score 1669.2; DB 9; Length 1708;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
QY	268	CCAGGCAAGCCATCATGAAGCTGAATGGCCACCACTGGAGAAACATCGCCCTGAAAGTCT	327						
DB	30	CCCCGGAGCCATCATGAAGCTGAATGGCCACCACTGGAGAAACATCGCCCTGAAAGTCT	89						
QY	328	CCTACATCCCAGATGAGCAGATAGCACAGGACCTGAGAAATGGGGCGCGAGGGGCTTTG	387						
DB	90	CCTACATCCCAGATGAGCAGATAGCACAGGACCTGAGAAATGGGGCGCGAGGGGCTTTG	149						
QY	388	GCTCTCGGGTACGCCCCGCGAGGCTACCTGTGGCAGCGGGGCCCCAGCAAGCAGC	447						
DB	150	GCTCTCGGGTACGCCCCGCGAGGCTACCTGTGGCAGCGGGGCCCCAGCAAGCAGC	209						
QY	448	AGCAAGTGGACATCCCCCTTCGGCTCTGTGTGCCCACTATGTGGGTGCCATTATTG	507						
DB	210	AGCAAGTGGACATCCCCCTTCGGCTCTGTGTGCCCACTATGTGGGTGCCATTATTG	269						
QY	508	GCAAGGAGGGGCCACCATCGCAACATCAAAAACAGCCAGTCCCAAGATAGACGTGC	567						
DB	270	GCAAGGAGGGGCCACCATCGCAACATCAAAAACAGCCAGTCCCAAGATAGACGTGC	329						
QY	568	ATAGAGAGGAAACGCGAGTGCAGTGAAGAGCCATCAGTGTGCCTCCACCTCGAGG	627						
DB	330	ATAGAGAGGAAACGCGAGTGCAGTGAAGAGCCATCAGTGTGCCTCCACCTCGAGG	389						
QY	628	GCTGCTCTCGCTTGAAGATGATCTTGGAGATTATGATAAAGAGGCTAAGGACACCA	687						
DB	390	GCTGCTCTCGCTTGAAGATGATCTTGGAGATTATGATAAAGAGGCTAAGGACACCA	449						
QY	688	AAACGGCTACGAGGTTCCCTGGAAGATCCTGGGCCATTAATACTTTGAGGGGCTTCA	747						
DB	450	AAACGGCTACGAGGTTCCCTGGAAGATCCTGGGCCATTAATACTTTGAGGGGCTTCA	509						
QY	748	TTGGCAAGGAGGACCGAACCTGAAAGAGGTAGACGAGATACCGAGCAAAAATCACCA	807						
DB	510	TTGGCAAGGAGGACCGAACCTGAAAGAGGTAGACGAGATACCGAGCAAAAATCACCA	569						
QY	808	TCTCTCTGTTGCAAGACCTTACCTTTACAAACCTTGAGAGGACCATCACTGTGAAGGGG	867						
DB	570	TCTCTCTGTTGCAAGACCTTACCTTTACAAACCTTGAGAGGACCATCACTGTGAAGGGG	629						
QY	868	CCATCGAGAAATTTGTCAGGGCCGAGCGAAATAATGAAGAAAGTTGCGGAGGCTATG	927						
DB	630	CCATCGAGAAATTTGTCAGGGCCGAGCGAAATAATGAAGAAAGTTGCGGAGGCTATG	689						
QY	928	AGAATGATGTGCTGCATGAGCTCTCAGCTATCCCTGGCCCTGACCTGGCTGCTGTAG	987						
DB	690	AGAATGATGTGCTGCATGAGCTCTCAGCTATCCCTGGCCCTGACCTGGCTGCTGTAG	749						
QY	988	GTCTTTTCCAGCTTCATCGAGGAGTCCCGCGGCTCCCGAGCAGCTTACTGGGGCTG	1047						
DB	750	GTCTTTTCCAGCTTCATCGAGGAGTCCCGCGGCTCCCGAGCAGCTTACTGGGGCTG	809						
QY	1048	CTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTTGAGAGGTGTTATCCCG	1107						
DB	810	CTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGTTGAGAGGTGTTATCCCG	869						
QY	1108	CCGAGCAGTGGGGCCATCATCGGCAAGAGGGGAGCAGATCAAAACAGCTCTCCCGGT	1167						
DB	870	CCGAGCAGTGGGGCCATCATCGGCAAGAGGGGAGCAGATCAAAACAGCTCTCCCGGT	929						

RESULT 3
US-09-873-637-1
; Sequence 1, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-873-637-1

QY	1168	TTGCCAGGGCTCCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTCGTATGG	1227
DB	930	TTGCCAGGGCTCCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTCGTATGG	989
QY	1228	TTATCATCACTGGACCGCCCAATTTCAAGGCTCAGGGAGAAATCTATGCAAAAC	1287
DB	990	TTATCATCACTGGACCGCCCAATTTCAAGGCTCAGGGAGAAATCTATGCAAAAC	1049
QY	1288	TCAAGGAGGAACTTTTGGTCCCAAGGAGGAAAGTGAAGTGGAGAGCCCATACCTG	1347
DB	1050	TCAAGGAGGAACTTTTGGTCCCAAGGAGGAAAGTGAAGTGGAGAGCCCATACCTG	1109
QY	1348	TGCCAGCATCAGCAGCTGGCGGCTCATTTGCAAGGTGGAAGAAACGCTGAACGAGTTGC	1407
DB	1110	TGCCAGCATCAGCAGCTGGCGGCTCATTTGCAAGGTGGAAGAAACGCTGAACGAGTTGC	1169
QY	1408	AGAAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACC	1467
DB	1170	AGAAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACC	1229
QY	1468	AGGTATCTGTGAAATCATTCGGAATTTCTATGCCAGTCAGATGGCTCAACGGAGATGCC	1527
DB	1230	AGGTATCTGTGAAATCATTCGGAATTTCTATGCCAGTCAGATGGCTCAACGGAGATGCC	1289
QY	1528	GAGCATCTCGCCAGGTTAAGCAGCAGCATTCAGAGGGACAGAGTAACACGAGGCCACAGG	1587
DB	1290	GAGCATCTCGCCAGGTTAAGCAGCAGCATTCAGAGGGACAGAGTAACACGAGGCCACAGG	1349
QY	1588	CACGGAGGAAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGGGACGAA	1647
DB	1350	CACGGAGGAAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGGGACGAA	1409
QY	1648	ATCGAGAGTGTCTCTCCCGGAGGCTGAGAAATGAGTGGGAAATCCGGGACACNTGGGC	1707
DB	1410	ATCGAGAGTGTCTCTCCCGGAGGCTGAGAAATGAGTGGGAAATCCGGGACACNTGGGC	1469
QY	1708	CGGCTGTAGATCAGGTTTGGCCACTTCATTCAGAAAGATGTTCCAGTGAGGAAACCTCGA	1767
DB	1470	CGGCTGTAGATCAGGTTTGGCCACTTCATTCAGAAAGATGTTCCAGTGAGGAAACCTCGA	1529
QY	1768	TCTNTCAGCCCCCAACACCCCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTCTAG	1827
DB	1530	TCTNTCAGCCCCCAACACCCCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTCTAG	1589
QY	1828	AAATNTAGCGCAAGGCACTTTTAAACGTGGATTTGTTTAAAGAGCTCTCCAGGCCCCAC	1887
DB	1590	AAATNTAGCGCAAGGCACTTTTAAACGTGGATTTGTTTAAAGAGCTCTCCAGGCCCCAC	1649
QY	1888	CAAGAGGTGGATCACCTCAGTCGGAAGAAAATAAAATTTCTTCAGGTTTAAAAA	1946
DB	1650	CAAGAGGTGGATCACCTCAGTCGGAAGAAAATAAAATTTCTTCAGGTTTAAAAA	1708

Db 61 CCAACAGAGCCGGGGTGGAGGGCGAGTGCTCAGCTTCCGGGTAGGAGCCGAAAA 120
Qy 121 TTCAAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCCCTGCTGGCTC 180
Db 121 TTCAAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCCCTGCTGGCTC 180
Qy 181 AGTATGTTACAGTACAGAGTCTGAGCAAGTGAACCCGAGAGTGAACCGAGAGTGGTGA 240
Db 181 AGTATGTTACAGTACAGAGTCTGAGCAAGTGAACCCGAGAGTGAACCGAGAGTGGTGA 240
Qy 241 ATGTCACTATTTCACACCGGAGAGCAGACCCAGCAAGCCATCATGAAGCTGAATGCCACC 300
Db 241 ATGTCACTATTTCACACCGGAGAGCAGACCCAGCAAGCCATCATGAAGCTGAATGCCACC 300
Qy 301 AGTTGGAGAACCATGCGCTGAGGTCTTCCTATCATCCCGATGAGCAGATACACAGGGAC 360
Db 301 AGTTGGAGAACCATGCGCTGAGGTCTTCCTATCATCCCGATGAGCAGATACACAGGGAC 360
Qy 361 CTGAGATGGCGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACCTG 420
Db 361 CTGAGATGGCGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCCAGGGCTCACCTG 420
Qy 421 TGGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCCTGGTGC 480
Db 421 TGGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCCCGCTTCGGCTCCTGGTGC 480
Qy 481 CCACCCAGTATGTTGGTGGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAAA 540
Db 481 CCACCCAGTATGTTGGTGGCCATTATTGGCAAGGAGGGGGCCACCATCCGCAACATCAAA 540
Qy 541 AACAGACCCAGTCCAGATAGACGTGCATAGGAGGAGAGCAGAGTGCAGCTCAAAAG 600
Db 541 AACAGACCCAGTCCAGATAGACGTGCATAGGAGGAGAGCAGAGTGCAGCTCAAAAG 600
Qy 601 CCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCGCTTGTAGATGATCTTGAGA 660
Db 601 CCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCCGCTTGTAGATGATCTTGAGA 660
Qy 661 TTATGATTAAGAGGCTTAAGGACACCAAAACCGGTGACAGGTTCCCTGAAGATCTGG 720
Db 661 TTATGATTAAGAGGCTTAAGGACACCAAAACCGGTGACAGGTTCCCTGAAGATCTGG 720
Qy 721 CCCATATAATTTGTAGGCGTCTCATTCGCAAGGAGAGCAGCAACCTCAAGAGGTAG 780
Db 721 CCCATATAATTTGTAGGCGTCTCATTCGCAAGGAGAGCAGCAACCTCAAGAGGTAG 780
Qy 781 AGCAAGATACCGAGACAAATAACCATCTCTCGTTGCAAGACCTTACCCTTTACAACC 840
Db 781 AGCAAGATACCGAGACAAATAACCATCTCTCGTTGCAAGACCTTACCCTTTACAACC 840
Qy 841 CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTGTCAGGGCGGAGCAGAAA 900
Db 841 CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTGTCAGGGCGGAGCAGAAA 900
Qy 901 TAATGAAGAAAGTTCCGGAGGCTTATGAGATGATGCTGGCTGCATGAGCTCTCACCTGA 960
Db 901 TAATGAAGAAAGTTCCGGAGGCTTATGAGATGATGCTGGCTGCATGAGCTCTCACCTGA 960
Qy 961 TCCCTGGCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGGCGAGTCCCGC 1020
Db 961 TCCCTGGCTGAACCTGGCTGTAGTCTTTTCCAGCTTCATCCAGGCGAGTCCCGC 1020
Qy 1021 CGCTCCAGCAGCGTACTGGGGCTGCTCCCTATAGCTCTTATGAGGCTCCGAGC 1080
Db 1021 CGCTCCAGCAGCGTACTGGGGCTGCTCCCTATAGCTCTTATGAGGCTCCGAGC 1080
Qy 1081 AGGAGATGTTGAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATCGCAAGAG 1140
Db 1081 AGGAGATGTTGAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATCGCAAGAG 1140
Qy 1141 GGCAGACATCAACAGCTCTCCCGTTTCCAGCGCCTCCATCAAGATTGACCAACCG 1200
Db 1141 GGCAGACATCAACAGCTCTCCCGTTTCCAGCGCCTCCATCAAGATTGACCAACCG 1200

Db 1141 GGCAGACATCAACAGCTCTCCCGTTTCCAGCGCCTCCATCAAGATTGACCAACCG 1200
Qy 1201 AAAACACCTGACTCCAAAAGTTTCGTATGTTATCATCACTGGACCGCCAGAGGCCCAATTCA 1260
Db 1201 AAAACACCTGACTCCAAAAGTTTCGTATGTTATCATCACTGGACCGCCAGAGGCCCAATTCA 1260
Qy 1261 AGGTCAGGGAAGAACTCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAG 1320
Db 1261 AGGTCAGGGAAGAACTCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAG 1320
Qy 1321 AAGTGAAGCTGGAGACCCACATACGTGTCAGGATCAGCAGCTGGCGGGTCTATTGGCA 1380
Db 1321 AAGTGAAGCTGGAGACCCACATACGTGTCAGGATCAGCAGCTGGCGGGTCTATTGGCA 1380
Qy 1381 AAGTGGAAAAAAGCGTGAACGAGTTGACAGGATTCAGCGCAGCTGAGGTGCTAGTACCAA 1440
Db 1381 AAGTGGAAAAAAGCGTGAACGAGTTGACAGGATTCAGCGCAGCTGAGGTGCTAGTACCAA 1440
Qy 1441 GAGACAGACCCCTGATGAGAAACGACAGCTCATCTGTAATAATCATCGGACATTTCTATG 1500
Db 1441 GAGACAGACCCCTGATGAGAAACGACAGCTCATCTGTAATAATCATCGGACATTTCTATG 1500
Qy 1501 CCAGTCAGATGGCTCAACGGAAGTCCGAGACATCTGGGCCAGGTTAAGCAGCAGCATC 1560
Db 1501 CCAGTCAGATGGCTCAACGGAAGTCCGAGACATCTGGGCCAGGTTAAGCAGCAGCATC 1560
Qy 1561 AGAAGGACAGAGTAACACGAGCCAGGCGCAGGAGGAAGTACCCAGCCCTCCCTGTCCT 1620
Db 1561 AGAAGGACAGAGTAACACGAGCCAGGCGCAGGAGGAAGTACCCAGCCCTCCCTGTCCT 1620
Qy 1621 TNGAGTCCAGGACAAACAAAGGGCAGAAATCGAGAGTGTGCTCTCCCGCAGGCTGAGA 1680
Db 1621 TNGAGTCCAGGACAAACAAAGGGCAGAAATCGAGAGTGTGCTCTCCCGCAGGCTGAGA 1680
Qy 1681 ATGAGTGGGAAATCCGGGACACACNTGGCGCTGTAGATCAGGTTTGGCCACTTGATGA 1740
Db 1681 ATGAGTGGGAAATCCGGGACACACNTGGCGCGCTGTAGATCAGGTTTGGCCACTTGATGA 1740
Qy 1741 GAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGGCCCAACACCCACCCAAATGGCCCC 1800
Db 1741 GAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGGCCCAACACCCACCCAAATGGCCCC 1800
Qy 1801 AACACTGTTGCCCTCGGGTGTGAGAAATTTAGCGAAGCACTTTAAACGTGGAT 1860
Db 1801 AACACTGTTGCCCTCGGGTGTGAGAAATTTAGCGAAGCACTTTAAACGTGGAT 1860
Qy 1861 TGTTTAAAGAGCTCTCCAGGGCCCAAGAGGTTGATCACACCTCAGTGGGAGAAA 1920
Db 1861 TGTTTAAAGAGCTCTCCAGGGCCCAAGAGGTTGATCACACCTCAGTGGGAGAAA 1920
Qy 1921 AATAAAATTTCTTCAAGTTTAAAA 1946
Db 1921 AATAAAATTTCTTCAAGTTTAAAA 1946

RESULT 2

US-09-899-651-5
; Sequence 5, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 24, 2004, 13:04:51 ; Search time 837.722 Seconds
(without alignments)
11355.361 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacggagggtggg.....atttccttcagggtttaaaa 1946

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3216467 seqs, 2444149694 residues
Total number of hits satisfying chosen parameters: 6432934

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1941	99.7	1946	9	US-09-899-651-7
2	1669.2	85.8	1708	9	US-09-899-651-5
3	1301	66.9	2224	9	US-09-873-637-1
4	1291.4	66.4	1738	16	US-10-062-674-2082
5	1085.4	55.8	1182	9	US-09-764-864-327
6	1069.8	55.0	1129	9	US-09-764-864-745
7	674.2	34.6	1740	9	US-09-735-705-347
8	674.2	34.6	1740	9	US-09-850-716A-347
9	674.2	34.6	1740	9	US-09-897-778-428
10	674.2	34.6	1740	13	US-10-007-700-347
11	674.2	34.6	1740	15	US-10-117-982-347
12	674.2	34.6	1740	15	US-10-117-982-478
13	674.2	34.6	1740	15	US-10-313-986-347
14	674.2	34.6	1740	16	US-10-313-986-478

15	674.2	34.6	1743	9	US-09-897-778-447
16	674.2	34.6	1743	9	US-09-897-778-450
17	674.2	34.6	1743	13	US-10-007-700-447
18	674.2	34.6	1743	13	US-10-007-700-450
19	674.2	34.6	1743	15	US-10-117-982-447
20	674.2	34.6	1743	15	US-10-117-982-450
21	674.2	34.6	1743	16	US-10-313-986-447
22	674.2	34.6	1743	16	US-10-313-986-450
23	674.2	34.6	1764	9	US-09-850-716A-428
24	674.2	34.6	1764	9	US-09-897-778-428
25	674.2	34.6	1764	13	US-10-007-700-428
26	674.2	34.6	1764	15	US-10-117-982-428
27	674.2	34.6	1764	16	US-10-313-986-428
28	674.2	34.6	1799	16	US-10-313-986-485
29	674.2	34.6	34555	15	US-10-117-982-479
30	674.2	34.6	34555	16	US-10-313-986-479
31	672.6	34.6	4159	9	US-09-899-651-4
32	672.6	34.6	4181	9	US-09-735-705-175
33	672.6	34.6	4181	9	US-09-954-456-715
34	672.6	34.6	4181	9	US-09-850-716A-175
35	672.6	34.6	4181	9	US-09-897-778-175
36	672.6	34.6	4181	10	US-09-466-396A-175
37	672.6	34.6	4181	13	US-10-007-700-175
38	672.6	34.6	4181	15	US-10-117-982-175
39	672.6	34.6	4181	16	US-10-313-986-175
40	672.6	34.6	4434	13	US-10-116-802-145
41	664.6	34.2	1740	15	US-10-117-982-483
42	664.6	34.2	1740	16	US-10-313-986-483
43	585.4	30.1	3283	9	US-09-899-651-8
44	585.4	30.1	3667	15	US-10-097-340-146
45	585.4	30.1	3667	17	US-10-648-593-48

ALIGNMENTS

RESULT 1
US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. US20020111470A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-7

Query Match 99.7%; Score 1941; DB 9; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCTGTAGCGAGGGCTGGGGGCTCTGTCCTTCTTGGCGGCTCGCGCTCAGC 60
DB 1 GCTGTAGCGAGGGCTGGGGGCTCTGTCCTTCTTGGCGGCTCGCGCTCAGC 60
QY 61 CCACCACGAGCGCGGGTGGAGGGGAGTGTCTCAGCTTCCCGGTTAGGAGCGGAAAA 120

This Page Blank (uspto)

Mon Jul 26 09:17:14 2004

Db 1683 GGTGAAGCAGCAGGAGCAGAAATACCCCTCAGGGAGTCGCCCTCACAGCGCCAGCAAGTGAGG 1742

QY 1604 AGCCC 1608

Db 1743 CTCCC 1747

Search completed: July 24, 2004, 16:48:13
Job time : 143.138 secs

Db 1680 AAAAATTGTTAGTCTTAAAGAGAGGTGAAAATTGAGGCTCATATCAGAGTGCCATCCT 1739
QY 1357 CAGCAGCTGGCCGGTCAATGGCAAGAGTGAAGAAACGGTGAACAGTTGAGAGTTTGA 1416
Db 1740 TTGCTGCTGGCAGAGTTATTGGAAAGAGGAGCAAAACGGTGAATGAACTTCAGAAATTGT 1799
QY 1417 CGGCAGCTGAGGTGGTGTAGTACCAAGAGACCAAGACCCCTGATGAGAACCAACAGGTCTATCG 1476
Db 1800 CAAGTGCAGAGTTGTTGTCCTCGTGACCAAGACACCTGATGAGAAATGACCAAGTGGTTG 1859
QY 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAAGTGGTCAACGGAAGATCCGACATCC 1536
Db 1860 TCAAAATAACTGGTCACTTCTATGCTGCCAGGTGCGCCAGAGAAAATTTCAGGAATTC 1919
QY 1537 TGGCCCAAGGTTAAGCAGCAGCATCAGAG 1565
Db 1920 TGAATCAGGTAAGCAGCAGCAACCAACAG 1948

RESULT 15

US-09-061-709-8
; Sequence 8, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-8

Query Match 30.1%; Score 585.4; DB 3; Length 3283;
Best Local Similarity 64.1%; Pred.No. 1.4e-160;
Matches 964; Conservative 0; Mismatches 481; Indels 60; Gaps 3;

QY 104 GGTAGAGCCGGAAATTCAAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACT 163
Db 303 GCTAAGGAGCAGGAAAATTCAGATTTCGAAACATCCCTCTCAGCTGCAGTGGGAGGTGT 362
QY 164 GGNACGCTGCTGGCTCAGTATGGTACAGTAGAATCTGTGACAAAGTGAACACCGAGAG 223
Db 363 GGATGGACTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAAGTCAACACAGACAC 422
QY 224 TGAGACGGCAGTGGTGAATGTCACTTATCCACCGGGAGCAGACCGAGCCAGCCATCAT 283
Db 423 AGNAACCGCGTGTCTCAAGTCAATATGCAACAGAGAGAAGAAAATAGCCATGGA 482
QY 284 GAAGTGAATGGCCACAGTTGGAGAACCATGCCCTGAAAGGTCTCTACATCCCGCATGA 343
Db 483 GAAGCTAAGCGGCGATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGATGA 542
QY 344 GCAGATAGCACAGGGACCTGAGAAATGGGCGCGAGGGGGCTTGGCTCTCGGGGTGAGCC 403
Db 543 -----AGAGGTGAGCTCCCTTTGGCCCTTCAGCGAGGCCAGCGTGGGGACCACTC 593
QY 404 CCGCCAGGGCTCACTGTGGCAGCGGGGGCCCGCAGCCAGCAGCAAGTGGACATCCC 463
Db 594 TTCCCGGAGCAGAGGCCACCGCCCTGGGGGCACTTCTCAGGGCCAGACGATTGATTTCCC 653
QY 464 CTTGGGCTCTCTGGTCCCAACCCAGATATGTGGGTGCCATTATTGGCAAGGAGGGGCCAC 523

Db 654 GCTCGGATCCTGTCCTCCACCCAGTTTGTGGTGCATCATCGGAAAGAGGGCTTGAC 713
QY 524 CATCCGCAACATCACAAAAAGAGACCCAGTCCAGATAGAGTGCATAGGAGGAGAACGC 583
Db 714 CATAAAGAACATCACTTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTC 773
QY 584 AGGTGAGCTGAAAGAACCCATCAGTGTGCACTCCACCCCTGAGGGGTGCTCTCCGCTTG 643
Db 774 TGGAGCTGCAGAGAAAGCCTGTCCATCCATCCACCCAGAGGGGACTTCTGAAGCATG 833
QY 644 TAAGATGATCTTGGAGATATGATTAAGAGGCTTAAGACACCAAAACGGCTGACGAGGT 703
Db 834 CCGCATGATCTTGAATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAAGAGAT 893
QY 704 TCCCTGAAAGATCCTGGCCCATATAACTTTGTAGGCGTCTCATTTGGCAAGGAAGACG 763
Db 894 TCCTCTGAAATCTTGGCACACATGGCTTGGTGGAGACTGATTGGAAGAAAGAGCAG 953
QY 764 GAACCTGAAGAGGTGAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGA 823
Db 954 AAATTTGAAGAAAATTGAACATGAAACAGGGACCAAGATAACAATCTCATCTTTGCAGGA 1013
QY 824 CTTTACCCTTTACAACCTGAGAGGACCATCACTGTGAAGGGGCCCATCGAGAAATGTTG 883
Db 1014 TTTGAGCATATACAACCCGAAAGAACCATCACTGTGAAGGGGCACAGTTGAGGCTGTGC 1073
QY 884 CAGGGCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCTATGAGATGATCTGGCTGC 943
Db 1074 CAGTGTGAGATGAGATTTATGAAGAGCTGCTGAGGCTTTGAAAATGATATGCTGGC 1133
QY 944 CATGAGCTCTACCTGATCCCTGGCTGAAACCTGGCTGCTGTAGGTCTTTTCCAGCTTC 1003
Db 1134 TGTTAACACCCACTCCGGATACTTCTCCAGCCTG-----TAC 1170
QY 1004 ATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTTGGGGCTGCTCCCTATAGTCTCTT 1063
Db 1171 CCCATCACAGTTTGGCCCGTTCCCGCATCATCTCTTA----- 1211
QY 1064 TATGAGGCTCCCGAGCAGGAGATGTCAGAGTGTATATCCCGCCGAGCAGTGGCGC 1123
Db 1212 -----TCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGCGC 1262
QY 1124 CATCATCGCAAGAGGGGAGCACAATCAACAGCTCTCCCGTTTGCAGAGCCCTCCAT 1183
Db 1263 CATCATCGGGAAGAGGGGGCACAATCAACAGCTGGGAGATTCCCGGAGCTCTAT 1322
QY 1184 CAAGATTGCACCAACCGGAAACCTGACTCCCAAGTTCCGTATGTTTATCATCTGACC 1243
Db 1323 CAAGATTGCCCTGCGGAAGGCCAGAGCTCAGCGAAAGATGGTTCATCATCCGGGCG 1382
QY 1244 GCCAGAGGCCAAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGAGAGAGAACTT 1303
Db 1383 ACCGGAAGCCAGTTCAAGGCCAGGACGATCTTTGGGAACCTGAAAGAGGAAACTT 1442
QY 1304 CTTTGTCCCAAGAGGAGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCAGC 1363
Db 1443 CTTTAAACCCCAAGAGAGTGAAGTGGAGCGCATATCAGAGTGCCCTCTCCACAGC 1502
QY 1364 TGGCCGGGTCTATTGGCAAGGTGGAAGAACGGTGAACGGTTGCAAGATTTGACCGCAGC 1423
Db 1503 TGGCCGGGTGATTGGCAAGGTGGCAAGCCCGTGAACCGAACTGCAAGACTTAAACAGTGC 1562
QY 1424 TGAGTGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACCGACAGGTCATCGTGAAT 1483
Db 1563 AGAAGTCATCGTGCCTCGTGACCAACCGCCAGATGAATGAGGAGTGCATCGTCAAT 1622
QY 1484 CATCGCAATTTCTATGCCAGTCAATGGCTCAAACGGAAGATCCGAGACATCTCGGCCA 1543
Db 1623 TATCGGGCACTTCTTTGTAGCAGACTGCACAGCGCAAGATCAGGGAATTTGTACAA 1682
QY 1544 GGTTAAGCAGCAGCATCAGAGGGGACAGATTAACAGGCCCCAGGACCGGAGAGGTGACC 1603

LOCATION: (3502)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3506)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3520)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3538)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3549)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3646)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3940)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3968)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3974)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4036)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4056)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4062)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4080)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4088)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4115)
OTHER INFORMATION: n=A,T,C or G

US-09-606-421B-175

Query Match 34.6%; Score 672.6; DB 4; Length 4181;
Best Local Similarity 66.6%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;
QY 103 GGGTTAGGAGCGGAAAATTCAAATCCGAAATATCCACCCAGCTCCGATGCGAAGTAC 162
DB 480 GCGAAAGGATTCGGAACCTTCAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGC 539
QY 163 TGGACAGCTGCTGGCTCAGTATGGTACAGTACGAGAACTGTGAGCAAGTGAACACCGAGA 222
DB 540 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTTCCAAACCGGGAGAGACCAAGGCCATCA 282

DB CGGAACTGCAGTTGTAAATGTAACCTATTCCAGTAAGGACCAAGCTAGACAAGCACTAG 659
QY TGAAGCTGAATGGCCACAGTTGGAGAACATGCCCTGAAGGTCTCTTACATCCCGATG 342
DB ACAAAGTGAATGGATTTCACTTGAAGAAATTTCACTTGAAGTACCTATATCCCTGATG 719
QY AGCAGATAGCACAGGACCTGAGAATGGGCGCGAGGGGCTTTGGCTCTCGGGCTCAGC 402
DB AATGGCGCGCCAGCAAAACCCCTTGACAGACGCCCGAGGTGCGGGGGCTTGGGAGA 779
QY CCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCGAGCCAGCAAGCAGCAGTGGACATCC 462
DB GGGGCTCTCAAGCAGGGGTCTCCAGATCCGTATCCAAGCAAAACCATGTGATTGC 839
QY CCTTTCGGCTCTCGTGGTCCCAACAGTATGGGTGCTTATTTGGGAGGAGGGGCGCA 522
DB CTCTGGCGCTGCTGGTTCCCAACCAATTTGTTGGAGCCATCATAGGAAAGAGGTGCCA 899
QY CCATCCGCAATCACAAACAGACCCAGTCCAGATAGAGTGCATAGAGGAGGAGAAAG 582
DB CCATTCGGAAACATCACCAACAGACCCAGTCTAAATCGATGTCCACCGTAAGAAATG 959
QY CAGGTGCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCGCTT 642
DB CGGGGCTGCTGAGAGTCACTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
QY GTAAGATGATCTTGGAGATTATGATAAAGAGGTAAAGACACCAAAACGGCTGACGAG 702
DB GTAAGTCTATTCTGGAGATTATGATAAAGAGGTCAAGATATAAAATTCACAGAGAGA 1079
QY TCCCTTGAAGATTTTACTCATATAAATTTGTTGGAGCTTCTCTCTCTCTCTCTCTCT 762
DB TCCCTTGAAGATTTTACTCATATAAATTTGTTGGAGCTTCTCTCTCTCTCTCTCTCT 1139
QY GGAACCTGAAGAGGTAGAGCAAGATACCGAGCAAAATACCAATCTCTCTCTCTCTCT 822
DB GAAATCTTAAATAAATTGAGCAAGACACAGACACTAAATACGATATCTCCATTGCGAG 1199
QY ACCTTACCTTTTAAACCTTGAGAGGACCATCCTGTGAGGGGGCCATCGAGATTTGT 882
DB AATTGACGCTGTATAATCCAGAACGCACTATTACAGTTTAAAGGCAATGTGAGACAT 1259
QY GCAGGCGGAGCAGGAATAATGAAGAAAGTTCCGGAGGCGCTATGAGATGATGTGGCTG 942
DB CCAAAGCTGAGGAGAGATCATGAAGAAATCAGGAGTCTTATGAATGATATTTGCTT 1319
QY CCATGA-----GCTCTCACCTGATCCCTGGCTGAAACCTGGCTGTGTAGTCTTTTCC 996
DB CTATGAATCTTCAAGCACTTTAATCTCTGGATTAAATCTGAACGCTTTGGGTCTGTTC 1379
QY CAGCTTTCATCCAGCGAGTCCCGCGCTCCCGAGAGCTTACTTGGGGCTGCTCCCTATA 1056
DB CACCCACTTTCAGGGATGCCACCTCCACCTCAGGGCGCCCTTCAGCCATGACTCTCT 1439
QY GCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCGCCAGGCGAG 1116
DB ACCCGCAGTTTGAGCAATCAGAAACGAGACGTTCATCAGTTTATCCAGCTCTATCAG 1499
QY TGGGGCCATCATCGCAAGAGGGGAGCAGCATCAAAACAGCTCTCTCCCGTTTGGCAGCG 1176
DB TCGGTGCCATCATCGCAAGCAGGGCCAGCAGCATCAAGCAGCTTTCTCGCTTTGCTGGAG 1559
QY CTCTCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTCGTAAGTTGTTATCATCA 1236
DB CTTCAATTAAGATTGCTCCAGCGGAAGCAGACATGCTAAAGTGAAGATGGTGATTATCA 1619
QY CTGGACCGCCAGGCGCCAAATCAAGGCTCAGGGAAGATCTATGCAAACTCAAGAGG 1296
DB CTGGACCCAGAGGCTCAGTTCAGGCTCAGGGAAGAAATTTATGAAATAATTAAGAAG 1679
QY AGAATCTTCTTGGTCCCAAGGAGAGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCAT 1356

```

; LOCATION: (4080)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A, T, C or G
US-09-542-615A-175

```

Query Match	34.6%	Score 672.6;	DB 4;	Length 4181;
Best Local Similarity	66.6%	Pred. No. 5.4e-186;		
Matches 979; Conservative	0;	Mismatches 484;	Indels 6;	Gaps 1;

[illegible]

Qy	997	CAGCTTCATCAGCGCAGTCCCGCGCTCCCCAGCAGGTTACTGGGGCTGCTCCCTATA	1056
Db	1380	CACCACTTCAGGGATGCCACTCCCACTCAGGGCCCCCTCAGCCATGACTCCTCCCT	1439
Qy	1057	GCTCCTTTATGCAGGCTCCGAGCAGAGATGGTGAGGTTTATCCCGCCGACGAGCAG	1116
Db	1440	ACCGCGAGTTTGAGCAATCAAGAACGGAGACTGTTCATCAGTTTATCCAGCTCTATCAG	1499
Qy	1117	TGGCGCGCATCATCGGCAAGAGGGGAGCACATCAACAGCTCTCCGGTTTGGCAGCG	1176
Db	1500	TCGCTGCCATATCGGCAAGCAGGGCCAGCACATCAAGCAGCTTCTCGCTTTCTGGAG	1559
Qy	1177	CTTCCATCAAGATTGCAACCAACCCGAAACACTCACTCAATCCAAAGTTCGTATGGTTATCATCA	1236
Db	1560	CTTCAATTAAAGATTGCTCCAGCGGAAGCACCATGCTAAAGTCAGATGGTATCATCA	1619
Qy	1237	CTGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAGAAATCTATGGCAAACTCAAGAGG	1296
Db	1620	CTGACCAACAGAGGCTCAGTTCAAGGCTCAGGGAGAAATTTATGAAAAAATTTAAAGAAG	1679
Qy	1297	AGAACTCTTTTGGTCCCAAGGAGAAATGAAGCTGGAGACCCATACAGTGTGCCAGCAT	1356
Db	1680	AAAACITTTTGTAGTCTTAAGAGAGGTGAAACTTCAGACTCATATCAGAGTCCATCCT	1739
Qy	1357	CAGCAGTGC CGGGTCA TTGGCAAAAGGTGGAAAAACGGTGAACGAGTTTGCAGAATTTGA	1416
Db	1740	TTGCTGTGTCAGAGTTATTGGAAAAAGGAGGCAAAAACGGTGAATGAACTTCAGAAATTTGT	1799
Qy	1417	CGGCAGCTAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGACGACCGAGGTCAATCG	1476
Db	1800	CAAGTCAGAAAGTTGTGTCCCTCGTGACCAAGACCTGATGAGAAATGACCAAGTGGTTG	1859
Qy	1477	TGAAAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAAGATCCGAGACATCC	1536
Db	1860	TCAAAATAACTGTCACCTTATGCTTTGCCAGTGTGCCCAGAGAAAAATTCAGGAAATTC	1919
Qy	1537	TGCCCCAGGTTAAGCAGCAGCATCAGAAG	1565
Db	1920	TGACTCAGGTTAAGCAGCAGCACCACACAG	1948

```

RESULT 14
US-09-606-421B-175
; Sequence 175, Application US/090606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure

```

1477 TGAAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
1860 TCAAAATAACTGGTCACTTCTATGCTTGCAGGTTGCCAGAGAAATTCAGGAAATTC 1919
1537 TGGCCCAAGGTTAAAGCAGCAGCATCAGAAG 1565
1920 TGACTAGGTAAGCAGCAGCAGCAACACAG 1948

RESULT 13
US-09-542-615A-175
; Sequence 175, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fangez, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3549)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A, T, C or G
; NAME/KEY: unsure

720 AAATGGCCGCCAGCAAAACCCCTTGACAGAGCCCGAGGTGCGCGGGGCTTGGGCGA 779
403 CCCGCCAGGCTCACCTGTGGCAGCGGGGCCCCAGCAAGCAGCAGCAAGTGGACATCC 462
780 GGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGATATCCAGCAGAAACCATGTGATTGC 839
463 CCCTTCGGCTTCCTGGTCCCAACCCAGTATGTTGGTGCATATTTGGCAAGAGGGGCCA 522
840 CTCTGGCCTGCTGTTCCCAACCAATTTGTTGGAGCCATCATAGGAAAGAGGTGCCA 899
523 CCATCCGCAATCATCAAAACAGACCAGTCCAAAGATAGAGTGCATAGGAAGAGAACG 582
900 CCATTCGGAACATCAACCAACAGACCAGTCTAAATTCGATGTCCACCGTAAGAAATG 959
583 CAGGTGAGCTGAAAGAGCATAGTGTGCACTCCACCCCTGAGGGTGTCTCCCTCGCTT 642
960 CGGGGCTGCTGAGAGTCAATTAATCTCTCTACTCTCTGAGGACCTCTCGGCTT 1019
643 GTAAGATGATCTTGGAGATTATGATAAAGAGGTAAGGACCAACAAACGGCTCAGCAGG 702
1020 GTAAGTCTATTCGAGATTATGCATAAAGGAAGCTCAAGATATAAAATTCACAGAAGA 1079
703 TTCCCTGAGATCCTCGGCCATATAATCTTTGAGGCGTCTCATTTGGCAAGAGGAC 762
1080 TCCCTTGAAGATTAGCTCATATAAATCTTTGTCAGCTCTTATTGTTAAAGAGGAA 1139
763 GGAACCTGAAGAGTGAAGCAAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAG 822
1140 GAAATCTTAAATAATTGAGCAAGACACAGACATTAATACGATATCTCAATTCGAGG 1199
823 ACCTTACCTTTTACAACCTCGAGAGGACCATCATCTGTGAAGGGGGCCATCGAGAAATGTT 882
1200 AATTGACGCTGATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGT 1259
883 GCAGGCGCGAGAGAAATATGAAGAAGTTCGGAGGCTTATCAGATGATGTTGGCTG 942
1260 CCAAAGCTGAGGAGGATCATGAAGAAATCAGGGGCTTTATGAAATGAAATGCTT 1319
943 CCATCA-----GCTCTCACCTGATCCCTGGCTCAACCTGGCTGCTGTAGTCTTTTCC 996
1320 CTATGAATCTCAAGCACATTAATTTCTGGAATTAATCTGNAACGCTTGGTCTGTTC 1379
997 CAGCTTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCCTATA 1056
1380 CACCACCTTCAGGATGCCACCTCCACTCAGGCGCCCTTCAGCCATGACTCTCCCT 1439
1057 GCTCTTTATGAGGCTCCGAGCAGGATGTTGTCAGGTGTTTATCCCGCCAGGAG 1116
1440 ACCCGAGTTTGAAGCAATCAGAAACGAGACTGTTTCATCAGTTTATCCAGCTCTATCAG 1499
1117 TGGGCGCATCATCGCAAGAGGGGCGAGCAGCATCAACAGCTCTCCCGTTTGGCAGG 1176
1500 TCGGTGCCATCATCGCAAGCAGGGCGAGCAGCATCAAGCAGCTTCTCGCTTGTGGAG 1559
1177 CTTCCATCAAGATTGCACCCAGCCCAACACCTGACTCCTAAAGTTCGATGTTATCATCA 1236
1560 CTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAAGATGTTGATTA 1619
1237 CTGGAACGCGAGGCGCAATTCAGGCTCAGGAGATCTTATGCAAACTCAAGGAGG 1296
1620 CTGGAACCAACAGGCTCAGTTCAAGGCTCAGGAGAGAAATTTATGAAATAAAGAG 1679
1297 AGAATCTCTTCTCCAGGAGGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCAT 1356
1680 AAAATCTTGTAGTCTTAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCATCCT 1739
1357 CAGCAGCTGCGCGGTTCATTGGCAAGGTGGAAAAACCGTGAACGAGTTGCAGAAATGA 1416
1740 TTGCTGCTGCGAGAGTTATTGGAAGAGGAGGCAAAACCGTGAATGAATTCAGAAATGT 1799
1417 CGCAGCTCAGGTGTTAGTACCAAGAGCAGACCCCTGATGAGAGCAGCAGGTCATCG 1476
1800 CAAAGTCAGAAGTTGTTGTTCCCTCGTGAACAGACACCTGATGAGATGACCAAGTGTTG 1859

```
QY 163 TGGACAGCCTGCTGGCTCAGTATGCTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGA 222
Db 540 TGGATAGTTTACTAGTCCAGTATGAGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGACGGCAGTGGTGAATGTCTACCTATTCCAAACCGGAGCAGACACAGGCAAGCCATCA 282
Db 600 CGGAAACTGCGAGTGTAAATGTAACTATTCCAGTAAGGACCAAGCTAGACAAGCACTAG 659
QY 283 TGAAGCTGAATGGCCACAGTGTGAGAACCAATGCCCTGAAGGTCTCTCTACATCCCGCATG 342
Db 660 ACAAACCTGAATGGATTTTCAGTTAGAGAAATTTCACTTTGAAAGTAGCCCTATATCCCTCATG 719
QY 343 AGCAGATAGCACAGGGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTACG 402
Db 720 AATGGCCGCGCCAGCAAAACCCCTTGCAGCAGCCCCAGGTCCGGGGGCTTGGGCGAGA 779
QY 403 CCGCCAGGGCTCACTGTGTGGCAGCGGGGCCCCAGCCAAAGCAGCAGCAAGTGGACATCC 462
Db 780 GGGGCTCCTCAAGGCAGGGGTCTCCAGGATCCGATCCGATCCAAGCAGAAACCATGTGATTGC 839
QY 463 CCTTCGGCTCCTGGTGCACCCAGTATGTGGTGCATATTATGGCAAGAGGGGCCA 522
Db 840 CTCTGGCTCTGGTGTCCACCCCAATTTGTGGAGCCATATAGGAAAGAGAGGTGCCA 899
QY 523 CCATCCGCAACATCACAAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGAGAAAG 582
Db 900 CCATTCGGAACATCAACAACAGCCAGTCTAAATCGATGTCCACCGTAAGAAATG 959
QY 583 CAGGTGAGCTGAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGTCTCTCGCTT 642
Db 960 CGGGGGTGTGAGAACTGATTAATCTCTCTACTCTGAAAGCACTCTGCGCTT 1019
QY 643 GTAAGATGATCTTGGAGATTATGCAAAAGAGCTAAGGACACCAAAACGGCTGACAGG 702
Db 1020 GTAAGTCTATCTGGAGATTATGCAAAAGAGCTCAAGATATAAAATTCACAGAGAGA 1079
QY 703 TTCCCTCAGAGATCTCTGGCCATAATAACTTTGTAGGGCTCTCATTTGGCAAGGAAGGAC 762
Db 1080 TCCCTTGAAGATTTTAGTCTAATAACTTTGTGGAGCTTTATTTGGTAAGAGGAA 1139
QY 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCAACATCTCTCTGTTCAAG 822
Db 1140 GAAATCTTAAATAAATTTGAGCAGACACACACTAAATCAAGATATCTCCATTGCGG 1199
QY 823 ACCTTACCTTTTAAACCTGTGAGAGCACTCACTGTGAGGGGCCATCGAGAAATGTT 882
Db 1200 AATTGAGCTGTATATCCAGAACGCACTATTACAGTTAAAGCAATGTGAGACATGT 1259
QY 883 GCAGGCGCAGGAGGAGAAATGAAGAAAGTTCCGGAGGCCTATGAGAAATGATGTGCTG 942
Db 1260 CCAAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAATGATATTGCTT 1319
QY 943 CCATGA-----GCTCTACCTGATCCCTGGCTGAACTGGCTGTGTAGGCTTTTCC 996
Db 1320 CTATGAATCTTCAAGCACATTTAAATCTCTGATTAAATCTGAACGCTTTGGGCTCTGTTCC 1379
QY 997 CAGCTTCATCCAGCGCAGTCCCGCCCTCCAGCAGCGTTACTGGGCTGTCTCCCTATA 1056
Db 1380 GACCCACTTCAGGATGCCACCTCCCACTCAGGGCCCCCTTCAGCCATGACTCTCCCT 1439
QY 1057 GCTCTCTTTATGAGGCTCCCGAGCAGAGATGGTGCAGGTGTTATCTCCCGCCAGGAG 1116
Db 1440 ACCCGAGTTTGAGCAATCAGAAACGAGACTGTTCATCAGTTTATCCAGCTCTATCAG 1499
QY 1117 TGGGCGCATCATCGCAAGAGGGGCGAGCATCAAAACAGTCTCTCCGGTTTGCAGCG 1176
Db 1500 TCGGTGCCATCATCGCAAGAGGGGCGAGCATCAAGCAGCTTCTCGCTTTGCTGGAG 1559
QY 1177 CTTCAATCAAGATTGCAACCAACCGAAACACCTGCACTCCAAAGTCTGTATGGTTATCATCA 1236
Db 1560 CTTCAATTAAGATTGCTCCAGCGGAAGCAGATGCTAAAGTGAGGATGGTGAATATCA 1619
QY 1237 CTGGACCGCCAGAGGCCCAATTCAGGGCTCAGGGAAGAAATCTATGSCAAACTCAAGGAG 1296
```

```
Db 1620 CTGACACCACAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGAAAAATTAAGAAG 1679
QY 1297 AGAATCTTCTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCAT 1356
Db 1680 AAACTTTGTTAGTCTTAAAGAGAGGTGAACTTGHAGCTCATATCAGAGTGGCATCTCT 1739
QY 1357 CAGCAGCTGCGCGGCTCAATTTGGCAAAAGGTGAAAAACGGTGAACGATTCAGAAATTTGA 1416
Db 1740 TTGCTGCTGCGCAGAGTATTGGAAAAAGGAGGCAAAACGGTGAATGAATTCAGAAATTTGT 1799
QY 1417 CGCAGCTGAGGTGGTAGTACCAAGAGACCCAGACCCCTGTATGAGAACGACCAAGTCTATCG 1476
Db 1800 CAAGTGCAGAAAGTGTGTGCTCCCTCGTGACCCAGACACCTGATGAGAATGACCAAGTGTG 1859
QY 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACGGAAGATCCGAGAGATCC 1536
Db 1860 TCAAAATACTGCTCACTTCTATGCTTGCAGGTTCGCCAGAAAAATTCAGAAAAATTC 1919
QY 1537 TGCCCGAGGTTAAGCAGCAGCATCAGAAAG 1565
Db 1920 TGACTCAGGTAAGCAGCAGCCCAACAACAG 1948

RESULT 12
US-09-480-884A-175
; Sequence 175, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(4181)
; OTHER INFORMATION: n=A,T,C or G
US-09-480-884A-175

Query Match 34.6%; Score 672.6; DB 4; Length 4181;
Best Local Similarity 66.6%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;

QY 103 GGGTTAGAGCCGGAAAAATTCAAATTCGGAATATTCCACCCAGCTCCGATCGGAGTAC 162
Db 480 GGCAAGGATTCGAAAACTTCAGATACGAAATATCCGCTCATTTACAGTGGAGGTGC 539
QY 163 TGGACAGCTGCTGCTCAGTATGCTACAGTAGAATCTGACGACAGTGAACACCGAGA 222
Db 540 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGACGGCAGTGGTGAATGTCACTATTCCAAACCGGAGCAGACCCAGGCAAGCATCA 282
Db 600 CGGAAACTGCAGTTGTTAAATGTAACTATTCCAGTTAAGCAACAGCTAGACAGCACTAG 659
QY 283 TGAAGCTGAATGGCCACCGAGTTGGAGAACCAATGCCCTGAAGGTCTCTACATCCCGCATG 342
Db 660 ACAAACCTGAATGGATTTTCAGTTAGAGAAATTTCACTTGAAGTAGCTCTATATCCCTCATG 719
QY 343 AGCAGATAGCACAGGACCTGAGAAATGGGCGCCGAGGGGCTTTGGCTCTCGGGGTACG 402
```

Db 1080 TCCCTTGAAGATTAGCTCATAATACTTTGTGGACGCTCTATTGGTAAGAAGAA 1139
QY 763 GGAACCTGAAGAGGTAGACAGATACCGAGACAAAAATCACCATCTCTCGTTGCAAG 822
Db 1140 GAATCTTAAATAAATTGAGCAAGACACAGACATAAATCACGATATCTCCATTCGAG 1199
QY 823 ACCTTACCTTTTACAACCTCGAGAGGACCATCACTGTGAAGGGGCCCATCGAGAAATTGTT 882
Db 1200 AATTGACGCTGTATAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1259
QY 883 GCAGGGCCGAGCGAATAATGAGAAAGTTCCGGAGGCCCTATGAGAAATGATGCTG 942
Db 1260 CCAAGCTGAGGAGAGATCATGAGAAATCAGGGAGTCTTATGAAATGATATGCTT 1319
QY 943 CCATGA-----GCTCTACCTGATCCCTGCGCTGAACCTGGCTGCTGTAGTCTTTTCC 996
Db 1320 CTATGAATCTCAAGACATTTAATCTCTGATTAATCTGAAACGCTTGGTCTGTGTC 1379
QY 997 CAGCTTCATCCAGGCGAGTCCCGCGCTCCAGAGCGGTTACTGGGGCTGCTCCCTATA 1056
Db 1380 CACCCACTTCAGGATGCCACTCCACCTCAGGGGCCCTTCAGCCATGACTCTCCCT 1439
QY 1057 GCTCCTTTATCAGGCTCCAGAGGAGATGGTGCAGTGTATATCCCGCCAGGCGAG 1116
Db 1440 ACCGCGAGTTTGAGCAATCAGAAACGAGACTGTTTCATCGTTTATCCAGCTCTATCAG 1499
QY 1117 TGGGCGCATCATCGGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGTTTGGCAGCG 1176
Db 1500 TCGGTGCCATCATCGGCAAGAGGGGCGAGCATCAACAGCTCTCTCGTTTCTGCTTGTGAG 1559
QY 1177 CTTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTTCTGATGTTATCATCA 1236
Db 1560 CTTCAATTAGATTCTCCAGCGAAGCACCAGATGCTAAAGTGAGGATGTTGATTATCA 1619
QY 1237 CTGACCCGCGAGGCCCAATTCAAGGCTCAGGAGAGATCTATGCGAACTCAGAGGAGG 1296
Db 1620 CTGGACCAACAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAAATAAAGAG 1679
QY 1297 AGAATCTTTTGGTCCCAAGAGGAGGAGTGAAGCTGGAGACCCACATAGCTGTCAGCAT 1356
Db 1680 AAAATCTTTGTAGTCTTAAAGAGAGGTTGAACCTTGAAGCTCATATCAGAGTGCATCT 1739
QY 1357 CAGCAGCTGGCGGGTCAATTGGCAAGGTGHAACCGGTGAACCGAGTTGCGAATTTGA 1416
Db 1740 TTGCTGCTGGCAGAGTTATTGGAAGAGGCGCAAAACGGTGAATGAATTCAGAAATTTG 1799
QY 1417 CGGCGAGCTGAGTGTAGTATCAAGAGACGACACCCCTGATGAGAACGACGAGTCAATCG 1476
Db 1800 CAAAGTGCAAGATTGTTGTCCTCGTGCACGACACCTGATGAGAAATGACCAAGTGGTTG 1859
QY 1477 TGAATATCATCGACATTTCTATGTCAGTGCAGTGCATGCTCAACGGAAGATCCGAGACATCC 1536
Db 1860 TCAAAATAACTGCTACCTTCTATGCTGTCAGGTTGTCAGGAGAAATTCAGGAAATTC 1919
QY 1537 TGGCCGAGGTTAAGCAGCAGCATCAGAAG 1565
Db 1920 TGACTCAGGTAAAGCAGCACCACACAG 1948

RESULT 11

US-09-643-597-175
; Sequence 175, Application US/09643597
; Patent No. 6426072

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun

Query Match 34.6%; Score 672.6; DB 4; Length 4181;

Best Local Similarity 66.6%; Pred. No. 5.4e-186;

Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;

QY 103 GGTTAGGAGCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTAC 162

Db 480 GGCAGAGGATTCGGAATTCAGATACGAAATATCCCGCTCATTTACGTGGAGGTGC 539

; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 175
; LENGTH: 4181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (3347)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3502)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3506)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3520)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3538)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3646)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3940)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3968)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (3974)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4036)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4056)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4062)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4080)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4088)
; OTHER INFORMATION: n=A,T,C or G
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G

US-09-643-597-175

523 CCATCCGCAACATCACAAACAGACCAGTCCAAAGATAGAGCTGCATAGGAAGAGAACG 582
|||||
900 CCATTCGGAACATCACAAACAGACCAGTCCAAATCGATGTCCACCGTAAGAAATG 959
583 CAGGTGACGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGCTGCTCCCTCGCTT 642
960 CGGGGGCTGCTGAGAAGTCGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019
643 GTAAGATGATCTTGAGATGATGATTAAGAGGCTTAAGGACACCAAAACGGCTGACGAG 702
1020 GTAAGTCTATTCTGGAGATTATGCAATAGGAAGCTCAAGATATAAAATTCACAGAAGA 1079
703 TTCCCTCTGAGATCTCGCCCATATTAATCTTGTAGGGCTCTCATTTGGCAGGAAGAC 762
1080 TCCCTTTGAGATTTAGCTCATATTAATCTTTGTGACGCTTAATTTGGTAAGAAGGAA 1139
763 GGAACCTGAAGAAGGTAGACGAATATCCGAGACAAATAATCACCATCTCTCTGTTGCAAG 822
1140 GAAATCTTAAAAAATTTGACCAAGACACAGACACTTAATCAGATATCTCCATTCGAGG 1199
823 ACCTTACCTTTACAAACCTTGAGAGGACCATCTGTTGAAGGGGCCCATCGAGAAATGTT 882
1200 AATTGACGCTGTAAATCCAGAACGCATATTAAGTTAAAGGCAATGTTGAGACATGTG 1259
883 GCAGGGCCGACGAGCAATTAATGAAGAAAGTTGCGGAGGCTATGAGAAATGATGCGCTG 942
1260 CCAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAATGATATTGCTT 1319
943 CCATGA-----GCTCTACCTGATCCCTGGCCCTGAACTGGCTGCTGTAGGCTTTTCC 996
1320 CTATGAATCTTCAAGCAATTAATCTCGGATTAATCTGAACGCTTGGGCTGTCTTCC 1379
997 CAGCTTCATCAGGCGAGTCCCGCGCTCCAGAGCGGTTACTGGGGCTGCTCCCTATA 1056
1380 CACCACTTCAGGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGACTCCTCCCT 1439
1057 GCTTCCTTTATGAGGCTCCCGAGCAGGAGATGTGCAAGGTGTTATCCCGCCCGCAGCAG 1116
1440 ACCGCGAGTTTGAGCAATCAGAAACGGAGACTGTTTATCATGTTATCCCGAGCTCTATCAG 1499
1117 TGGGCGCATATCTGGCAAGAGGGGAGCAGACATCAAAACAGCTCTCCCGGTTTGCAGCG 1176
1500 TCGGTGCCATCATCGGCAAGCGGCGCAGCACATCAAGCAGCTTCTCGCTTTGCTGGAG 1559
1177 CTTCCATCAAGATTGCAACACCGAACACCTGACTCCAAAGTTCGTAAGTTATCATCA 1236
1560 CTTCAATTAAGATTGCTCCAGCGGAAGCAGCAGATGCTAAAGTGAGGATGGTGAATTATCA 1619
1237 CTGGACCGCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGGAGG 1296
1620 CTGGACCCAGAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGGAATAATTAAGAAG 1679
1297 AGAATCTTTTGTGTCGAAGGAGAGTGAAGCTGGAGACCCACATACGTTGTCAGCAT 1356
1680 AAAACTTTGTTAGTCTTAAAGAGAGGTGAAACTTGAAGCTCATATCAGATGCCATCCT 1739
1357 CAGCAGCTCGCGGTCATTTGCAAGGTGGAAGAACCGTGAACGAGTTGACAGATTTGA 1416
1740 TTGCTCTGGCAGAGTTATGGAAGAGGAGGCAAAACGGTGAATGAATTCAGAAATTTGT 1799
1417 CGGCACTGAGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACCGAGTTCAG 1476
1800 CAAATGTCAGAGTTGTTGTCCTCTGTCGACGACGACCTGATGAGATGACCAAGTGGTG 1859
1477 TGAATATCATCGACATTTCTATGTCAGTCAAGTGGCTCAACGGAAGATCCGAGACATCC 1536
1860 TCAAAATAAATGCTGCTCTATGCTTGCAGGTTGCCAGGTTGCCAGAGAAAATTCAGGAAATTC 1919
1537 TGGCCAGGTTAAGCAGCAGCATCAAG 1565
1920 TGAATCAGGTAAGCAGCAGCAACACAG 1948

RESULT 10

US-09-899-651-4
; Sequence 4, Application US/09899651
; Patent No. 6576756

GENERAL INFORMATION:

; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; FILE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-4

Query Match 34.6%; Score 672.6; DB 4; Length 4159;
Best Local Similarity 56.8%; Pred. No. 5.4e-186;
Matches 979; Conservative 0; Mismatches 484; Indels 6; Gaps 1;

QY 103 GGTTAGGAGCCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGAAAGTAC 162
Db |||||
QY 480 GCAAGAGATTGCGAACTTCAGATACGAAATATCCGCTCATTTACAGTGGAGGTGC 539
Db |||||
QY 163 TGAACAGCTGCTGGCTCAGTATGGTACAGTACAGAACTGTGAGCAAGTGAACACCGAGA 222
Db |||||
QY 540 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACT 599
QY 223 GTGAGACGCGAGTGTGAATGTACCTATTCACCTATTCACCGGAGAGACAGCAGCAAGCATCA 282
Db |||||
QY 600 CGAAACTCAGTGTAAATGTAACTATTCAGTAAAGCAACCAAGCTAGACAGCACTAG 659
QY 283 TGAAGCTGAATGGCCACCACTGTTGGAAACCATGCTCCCTGAAGTCTCTACATCCCGCATG 342
Db |||||
QY 660 ACAACTGAATGATTTAGTATAGAAATTTTCACTTTGAAGTAGCTATATCCCTGATG 719
QY 343 AGCAGATACACAGGACCTGAGATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGACG 402
Db |||||
QY 720 AAATGGCCCGCCAGCAAAAACCCCTTTGACAGCCCGAGGTGCGCCGGGGCTTTGGGCAGA 779
QY 403 CCGCCAGGGCTCACCTGTGGCAGGGGGGGCCCGCCAGCAAGCAGCAGCAAGTGGACATCC 462
Db |||||
QY 780 GGGGCTCTCAGGCGAGGGGTCTCCAGGATCGTATCCAAGCAGAAACCATGTGATTTCG 839
QY 463 CCCTTCGGCTCCTGTGGTCCCACTGAGTGTGGGTGCCATTTATGGCAAGGAGGGGCCA 522
Db |||||
QY 840 CTCTGCGCTGCTGGTTCCTCCCACTTTGTTGGAGCCATCATAGGAAAGAGGTGCCA 899
QY 523 CCATCGCNAATCACAACAGACCCAGTCCAGATAGACGTGATAGGAGGAGAACG 582
Db |||||
QY 900 CCAATTCGAACATACCAACAGACCCAGTCTTAAATCGATGTCCACCGCTTAAAGAAATG 959
QY 583 CAGGTGACGTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGCTGCTCTCGCTT 642
Db |||||
QY 960 CGGGGCTGCTGAGAGTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGGGCTT 1019
QY 643 GTAAGATGATCTTGAGATTTATGCAATAAGAGGCTTAAGGACACCAAAACGGCTGACGAG 702
Db |||||
QY 1020 GTAAGTCTATTCTGAGATTTATGCAATAGGAAGCTCAAGATATAAAATTCAGAGAGA 1079
QY 703 TTCCCTCAGATCTCGGCCCATATAAATCTTTGTAGGGCTCTCTCANTGGCAAGGAGAC 762

	Matches	980;	Conservative	0;	Mismatches	483;	Indels	6;	Gaps	1;
QY	103	GGTTTAGGAGCGGAAAAATTC	AAATTC	CAAAATATTC	CAACCCAGCTCCGATGGGAAGTAC	162				
Db	230	GGCAAAGGATTCGGAACCTT	CAGATAC	GAATATCCCGCTCATTT	CAGTGGGAGGTGC	289				
QY	163	TGGACAGCCTGCTGGCTCAGT	TGGTAC	AGTAGAGAACTGTG	AGCAAGTGTAGAACCCGAGA	222				
Db	290	TGGTAGTTTACTAGTCAGT	TATCGAGT	TGGTGGAGAGCTGTG	AGCAAGTGAACACTGACT	349				
QY	223	GTGAGACGGCAGTGGTGAAT	CTCACCTT	ATCCACCGGGAGCAGAC	CAGGCAAGCCATCA	282				
Db	350	CGGAAACTGACAGTTGTTAA	TGTAACTTATCC	AGTAAGGACCAAGCTAG	CAACAGCACTAG	409				
QY	283	TGAAGCTGAATGGGCCAC	CCAGTTG	GAGAACCATG	CCCTGAAGGTCTCTACATCCCGCATG	342				
Db	410	ACAACTGAATGGATTCAGT	TAGAGAAATTTCA	CTTGAAGTAGCCTTAT	TATCCCTGATG	469				
QY	343	AGCAGATAGCACAGGACCT	GAGAACTGGG	CGCCGAGGGGGCTTTGG	CTCTCGGGGTGAGC	402				
Db	470	AAACGGCGCCACGACAAA	ACCCCTTG	CAGCAGCCCGAGGTG	CGCGGGGCTTTGGGCGA	529				
QY	403	CCCGCCAGGGCTCACCTGT	GCAGCGGGGGCCCCAG	CAAGCAGCAGCAAGTGG	ACATCC	462				
Db	530	GGGGCTCTCAAGCAGGGGT	CTCCAGATCCCGTAT	CCAAAGCAAAACCAATGT	GAATTTGC	589				
QY	463	CCCTTCGGCTCTGGTGCC	CAACCCAGTATGTGGT	GGCCATATTGGCAAGAGGGGGCCA	522					
Db	590	CTCTGCGCTGCTGGTTT	CCCCACCAATTTGTTG	AGCCATCATAGNAAAGAGGTGCCA	649					
QY	523	CCATCCGCAATCACAAA	CAGACCCAGTCC	CAAGTAGACGTGCAT	TAGGAAGGAGAACG	582				
Db	650	CCATTCGGAAATCATCC	AAAACAGACCCAGTCT	AAAAATCGATGTCC	ACCGTAAAGAAAATG	709				
QY	583	CAGGTGCAGCTGAAAAG	CCCATCAGTGTGC	ACTCCACCCCTGAGGGCTGCT	CTCCGCTT	642				
Db	710	CGGGGCTGCTGAGAGT	CGAATCTAATCTCT	CTACTCTCTGAAGGCACCTCT	GTGGGCTT	769				
QY	643	GTAAATGATCTTTGGAG	ATTATGCA	TAAAGAGGCTAAGGACAC	CAAAAACGGCTGACGAGG	702				
Db	770	GTAAGTCTAATCTGGAG	ATTATGCA	TAAAGAGGCTC	GAATATAAATTCACAGAGAGA	829				
QY	703	TTCCCTTGAAGATCTG	GGCCATAATAC	TTTGTAGGGCGTCTCAT	TGGCAAGAAAGGAC	762				
Db	830	TCCCTCTGGAATTTAG	CTCATATAACTTTCT	TGGACGCTCTATT	TGGTAAAGAAAGGAA	889				
QY	763	GGAACCTGGAAGAGGT	TAGACCAAGATAC	CAGACAAATATCAC	CAATCTCCTCGTTGCAAG	822				
Db	890	GAAATCTTAAAAAA	TTGACCAAGAC	CAGACACTAAATC	ACGATATCTCCATTCGAG	949				
QY	823	ACCTTACCCCTTTACA	CCCTGAGAGGACCAT	CACTGTGAAGGGGGCCAT	CGAATTTGTT	882				
Db	950	AAATTGACGCTGTAT	ATCCAGAAACGCACT	TATTACAGTTAAAGCAATGTT	TGACACATGTG	1009				
QY	883	GCAGGCCGACGAGGAA	TAATGCAAGAAAGTT	CGGAGGCCCTTAGA	AGATGATGGCTG	942				
Db	1010	CCAAAGCTGAGGAGG	AGATCATGAAGAAAT	CAGGAGTCTTATG	AAAAATGATTTGCTT	1069				
QY	943	CCATGA-----	GCTCTCACTCAT	CCCTGGCCCTGA	ACCTGGCTGTGATGGCTTTTCC	996				
Db	1070	CTATGATCTTCAAG	CACATTTAAATCT	TGGATTAATCTGA	ACGCTTTGGGCTGTGCTCC	1129				
QY	997	CAGCTTTCATCCAG	CGCAGTCCCGCCGCT	CCACAGCAGCGTTACT	TGGGGCTGTCTCCCTATA	1056				
Db	1130	CACCCACTTCAGG	ATGCCACCTCC	CACTCAGGGCCCCCTTC	CAGCCATGACTCTCCCT	1189				
QY	1057	GCTCTCTTATG	CAGGCTCCGAGC	AGAGATGGTG	CAGGTGTTATPCCCGCCCGCAG	1116				
Db	1190	ACCCGACAGTTG	AGCAATCAGAAAC	CGGAGCTGTT	CATCTGTTTATCCAGCTCTATCAG	1249				
QY	1117	TGGGCGCCATCAT	CGGCAGNAGGGG	CAGCAGATCAAA	CAGCTCTCCCGGTTGCGAGG	1176				
Db	1250	TGGTGCCATCAT	TCGCAAGCAGGGG	CAGCAGATCAAG	CAGCTTTCTCGCTTTGTTGAG	1309				

RESULT	8
US-09-6060	: Sequence
:	: Patent
:	: GENERAL
:	: APPLI
:	: APPLI
:	: APPLI
:	: APPLI
:	: APPLI
:	: APPLI
:	: APPLI
:	: APPLI
:	: TITLE
:	: FILE #
:	: CURRENT
:	: CURRENT
:	: NUMBER
:	: SOFTWARE
:	: SEQ ID
:	: LENGTH
:	: TYPE
:	: ORGAN
US-09-6060	

	Query Match	34.6%	Score 674.2	DB 4	Length 1740
	Best Local Similarity	66.7%	Pred. No. 1.1e-186		
	Matches 980	Conservative 0	Mismatches 483	Indels 6	Gaps 1
QY	103	GGGTTAGGAGCGGAAAATTCAAATCGAAATATTCACCCAGCTCCGATGGGAAGTAC	162		
Db	230	GGCAAGGATTGGAAACTTCAGATAGAAATATCCGCTCATTTACAGTGGGAGGTGC	289		
QY	163	TGGACAGCCTGCTGGCTCAGTATGTTACAGTAGAGAACTCTGAGCAAGTGAAACACCGAGA	222		
Db	290	TGGATATGTTTACTAGTCCAGTATGGAGTGGTGAGAGCTGTGAGCAAGTGAAACACTGACT	349		
QY	223	GTGAGACGCAGTGGTGAATGTCACTATTCCACCGGAGGAGACACGAGCAAGCCATCA	282		
Db	350	CGGAAACTCGAGTGTGAATGTGAACCTATTTCAGTAAAGGACCAAGCTAGACAGCACTAG	409		

Db	1010	CCAAGCTGAGGAGGAGATCATGAAGAAAAATCAGGAGCTCTATTGAAATGATTTGCTT	1069
QY	943	CCATCA-----GCTCTCACCTGATCCCTGCGCTGAACCTGGCTGCTAGGTCTTTTCC	996
Db	1070	CTATGAATCTTCAAGCACTTAATCTCTGATTAATCTGAACGCCTGGTCTGTTCC	1129
QY	997	CAGCTTTCATCCAGCGCAGTCCCGCGCCTCCAGCAGGCTTACCTGGGCTGCTCCCTATA	1056
Db	1130	CACCACCTTCAGGGATGCCACTCCCACTCAGGGCCCCCTTCAGCCATGACTCTCCCT	1189
QY	1057	GCTCCTTTATGAGCTCCCGAGCAGGATGCTGAGGTGTTTATCCCGCCAGGAG	1116
Db	1190	ACCGCAGTTTGAGCAATCAGAAAACGGAGCTGTTCTATCTGTTATCCAGCTCTATCAG	1249
QY	1117	TGGCGCCATCATCGGCAAGAGGGGAGCAGCATCAAAACAGCTCTCCCGTTTGCAGCG	1176
Db	1250	TCGGTGCCATCATCGGCAAGCAGGGCCAGCACAACAGCAGCTTCTCGCTTGTGGAG	1309
QY	1177	CCTCCATCAAGATTGACACCCCGAAACACCTGACTCCAAAGTTGCTATGTTTATCATCA	1233
Db	1310	CTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTAAAGTGAGGATGGTATTATCA	1369
QY	1237	CTGAGCCGAGAGGCCCAATTCAGAGCTCAGGGAAGATCTATGGCAAACTCAAGGAGG	1296
Db	1370	CTGGACCAACAGAGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGAAAAATTAAGAAG	1429
QY	1297	AGAACTTTTGGTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACCTGTGCCAGCAT	1356
Db	1430	AAAACTTTTGGTCTTAAAGAGGTTGAACTTGAAGCTCATATCAGAGTGCCATCCT	1489
QY	1357	CAGCAGCTGCGCGGCTCATTTGGCAAGGTGGAAAAACGGTGAACGAGTTGCAGAAATTGA	1416
Db	1490	TTGCTGCTGCAGAGATTATTGGAAAAAGGAGGCAAAACGGTGAATGAATTCAGAAATTGT	1549
QY	1417	CGGAGCTGAGTGGTAGTACCAAGAGACCAACCCCTCATGAGACGACGAGTCTATCG	1476
Db	1550	CAAGTGCAGAGTTGTTGCTCTGTCGACCAACCTCATGAGATGACCAAGTGGTTG	1609
QY	1477	TGAAATCATCGGACATTTCTATGCGAGTCAGATGGCTCAACCGAAGATCCGAGACATCC	1536
Db	1610	TCAAAATACTGTCATCTTATGTTGCGAGTTGCCAGAGAAAAATTCAGGAATTC	1669
QY	1537	TGCGCCAGGTTAAGCAGCAGCATCAGAAG	1565
Db	1670	TGACTCAGGTAAAGCAGCAGCAACACAG	1698
<p>RESULT 7 US-09-542-615A-347 ; Sequence 347, Application US/09542615A ; Patent No. 6518256 ; GENERAL INFORMATION: ; APPLICANT: Wang, Tongtong ; APPLICANT: Fan, Liqun ; APPLICANT: Kalos, Michael D. ; APPLICANT: Bangur, Chaitanya S. ; APPLICANT: Hosken, Nancy A. ; APPLICANT: Fanger, Gary R. ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY ; FILE REFERENCE: 210121.455C8 ; CURRENT APPLICATION NUMBER: US/09/542, 615A ; CURRENT FILING DATE: 2000-04-14 ; NUMBER OF SEQ ID NOS: 350 ; SOFTWARE: FastSeq for Windows Version 3.0 ; SEQ ID NO 347 ; LENGTH: 1740 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-542-615A-347</p>			
QY	942	CGAGGCGCAGAGGAAATATGAAGAAGTTGGGAGCCCTATCAGATCATGTGGCTG	942
Db	950	AATTCAGCTGTATTAATCCAGAACGCACTATTACAGTTAAAGGCAATTTGAGACATGTG	1009
QY	823	ACCTTACCTTTTACAACTCAGAGGACCACTCATGTGAAGGGGGCCATCGAGAAATGTT	882
Db	890	GAAATCTTAAAAAATTGAGCAAGACACAGACATTAATACCGATATCTCCATTCGAGG	949
QY	822	GGACCTCAAGAGTGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTGCAAG	882
Db	830	TCCGCTTGAAGATTATAGCTCATATAACTTTGTTGACGCTTATTGTTGAAGAGAA	889
QY	763	GGAACCTCAAGAGTGAGCAAGATACCGAGACAAAAATCACCATCTCTCTGTTGCAAG	822
Db	770	GTAAGTCTATTCTGGAGATTATGCATAAAGAGGCTAAAGCACCAAAACGGCTGACGAGG	829
QY	703	TTCCCTGGAAGATCCTGGCCCAATAACTTTGTTAGGGCTCTCATTTGGCAAGAGGAC	762
Db	643	GTAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAAGCACCAAAACGGCTGACGAGG	702
QY	642	CAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTT	642
Db	650	CCATTCGGAACATCACCAACAGACCCAGCTCAAAATCGATGTCACCCGTAAGAATAATG	709
QY	583	CAGTGCAGCTGAAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTT	642
Db	530	GGGGCTCCTCAAGCAGGGTCTCCAGATCCGATATCCAGACGAAACCATGTGATTTC	589
QY	463	CCCTTCGGCTCCTGGTCCCAACCCAGTATGCGGTGCAATTTGGCAAGAGGAGGCGCA	522
Db	590	CTCTGCGCTCTGTTTCCCAACCAATTTGTTGGAGCCATCATAGGAAAAAGAGGTGCCA	649
QY	523	CCATTCGCAACATCAAAAACAGACCCAGTCCAGATAGACGTCATAGGAAGGAGAACG	582
Db	530	GGGGCTCCTCAAGCAGGGTCTCCAGATCCGATATCCAGACGAAACCATGTGATTTC	589
QY	402	AGCAGATAGCACAGGACCTGAGATGGGCGCCGAGGGCTTTGCTCTCGGGGTGAGC	402
Db	470	AAACGGCGCCACAGCAAAACCCCTTCGAGCAGCCCCGAGGTCGCGGGGCTTTGGGAGA	529
QY	462	CCGCGCAGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAGTGGACATCC	462
Db	343	AGCAGATAGCACAGGACCTGAGATGGGCGCCGAGGGCTTTGCTCTCGGGGTGAGC	402
QY	342	TGAAGCTGAATGGCCACCACTTGGAGAACCATGCCCTGAAGGTCTCTCATATCCCCGATG	342
Db	410	ACAACTGTAATGATTTCACTTAGAGAAATTCACCTTGAAGTAGCTTATCCCTGATG	469
QY	283	TGAAGCTGAATGGCCACCACTTGGAGAACCATGCCCTGAAGGTCTCTCATATCCCCGATG	342
Db	350	CGGAACTGCAATGATTTCACTTAGAGAAATTCACCTTGAAGTAGCTTATCCCTGATG	469
QY	282	GTGAGACGGCAGTGGTGATGTCACTATTCCAAACGGGAGCAGACCCAGGCAAGCCATCA	282
Db	290	TGGATAGTTTACTAGTCCAGTATGAGTGGTGGAGAGCTGTGAGCAAGTGAACTGACT	349
QY	222	TGAGACGCTGCTGCTCAGTATGCTACAGTACAGAACTGTGAGCAAGTGAACTGACT	222
Db	230	GGCAAGGATTCGAAACTTCAGATAGAAATATCCGCTCATTTACAGTGGGAGTGC	289
QY	163	TGAGACGCTGCTGCTCAGTATGCTACAGTACAGAACTGTGAGCAAGTGAACTGACT	

;; PRIOR FILING DATE: 1998-04-17
;; NUMBER OF SEQ ID NOS: 8
;; SEQ ID NO 5
;; LENGTH: 1708
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
US-09-899-651-5

Query Match 85.8%; Score 1669.2; DB 4; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGSCCTCGAAGGTCT 327
Db 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACCAAGTTGGAGAACCATGSCCTCGAAGGTCT 89
328 CCTACATCCCGATGAGCAGATAGCAGAGCACTGAGAAATGGGGCCGAGGGGCTTTG 387
Db 90 CCTACATCCCGATGAGCAGATAGCAGAGCACTGAGAAATGGGGCCGAGGGGCTTTG 149
388 GCTCTCGGGGTACGCCCGCCAGGGCTCACCTGTGGCGAGCGGGCCCGCCAGCAAGCAGC 447
Db 150 GCTCTCGGGGTACGCCCGCCAGGGCTCACCTGTGGCGAGCGGGCCCGCCAGCAAGCAGC 209
448 AGCAAGTGGACATCCCCCTTCGGGTCTCTGTGTGCCCAACCCAGTATGTGGGTGCCATATTG 507
Db 210 AGCAAGTGGACATCCCCCTTCGGGTCTCTGTGTGCCCAACCCAGTATGTGGGTGCCATATTG 269
508 GCAAGAGGGGGCCACCTCCGGAACATCACAACACAGCCAGTCCCAAGTAGAGCTGC 567
Db 270 GCAAGAGGGGGCCACCTCCGGAACATCACAACACAGCCAGTCCCAAGTAGAGCTGC 329
568 ATAGMAGGAGACGCGAGTGCAGCTGAAAGCCATCAGTGTGCACTCCACCCCTGAGG 627
Db 330 ATAGMAGGAGACGCGAGTGCAGCTGAAAGCCATCAGTGTGCACTCCACCCCTGAGG 389
628 GCTGCTCTCCGCTGTGAAGATGATCTTGGAGATTATGATATAAGAGGCTAAGGACACCA 687
Db 390 GCTGCTCTCCGCTGTGAAGATGATCTTGGAGATTATGATATAAGAGGCTAAGGACACCA 449
688 AAACGGCTGACAGAGTTCCCTGAAAGATCTTGGCCCATTAATTAACCTTTAGGGGCTCTCA 747
Db 450 AAACGGCTGACAGAGTTCCCTGAAAGATCTTGGCCCATTAATTAACCTTTAGGGGCTCTCA 509
748 TTGGCAAGAGGAGCGAACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCA 807
Db 510 TTGGCAAGAGGAGCGAACCTGAAGAGGTAGACAGATACCGAGACAAATAATCACCA 569
808 TCTCTCTCGTTGCAAGACCTTACCTTTTACAAACCTTGAGAGGACCATCACTGTGAAGGGG 867
Db 570 TCTCTCTCGTTGCAAGACCTTACCTTTTACAAACCTTGAGAGGACCATCACTGTGAAGGGG 629
868 CCATCGAATTTGTCAGGGCCGAGAGGAATTAATGAAGAGTTTCGGGAGGCTATG 927
Db 630 CCATCGAATTTGTCAGGGCCGAGAGGAATTAATGAAGAGTTTCGGGAGGCTATG 689
928 AGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987
Db 690 AGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 749
988 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCCGAGCAGCGTTACTGGGGCTG 1047
Db 750 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCCGAGCAGCGTTACTGGGGCTG 809
1048 CTCCTTATAGCTCTTTATGAGGCTCCCGAGCAGGAGTGTGCGAGTGTATATCCCGG 1107
Db 810 CTCCTTATAGCTCTTTATGAGGCTCCCGAGCAGGAGTGTGCGAGTGTATATCCCGG 869
1108 CCAGGAGCTGGGCGCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGT 1167
Db 870 CCAGGAGCTGGGCGCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGT 929

1168 TTGCCAGGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGCTATGG 1227
Db 930 TTGCCAGGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGCTATGG 989
1228 TTATCATCACTGGACCGCCAGAGGCCCAATTCAGAGCTCAGGGAAGATTCATGGAACAAC 1287
Db 990 TTATCATCACTGGACCGCCAGAGGCCCAATTCAGAGCTCAGGGAAGATTCATGGAACAAC 1049
1288 TCAAGAGGAGAACTTCTTTGGTCCCAAGGGAAGTGAAGCTGGAGAGCCACATACCTG 1347
Db 1050 TCAAGAGGAGAACTTCTTTGGTCCCAAGGGAAGTGAAGCTGGAGAGCCACATACCTG 1109
1348 TGCCAGCATCAGCAGCTGGCCGGGTCAATGTGCAAAAGTGGAAAAACCGTGAACGAGTTGC 1407
Db 1110 TGCCAGCATCAGCAGCTGGCCGGGTCAATGTGCAAAAGTGGAAAAACCGTGAACGAGTTGC 1169
1408 AGAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACC 1467
Db 1170 AGAATTTGACGGCAGCTGAGGTGTAGTACCAAGAGACCAACCCCTGATGAGAACGACC 1229
1468 AGGTCACTGTGAAATCATTCGGAATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1527
Db 1230 AGGTCACTGTGAAATCATTCGGAATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1289
1528 GAGACATCTGSCCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACCAAGGCGCAGG 1587
Db 1290 GAGACATCTGSCCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAACCAAGGCGCAGG 1349
1588 CACGAGGAGAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACCAACACGCGGACAG 1647
Db 1350 CACGAGGAGAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACCAACACGCGGACAG 1409
1648 ATCGAGAGTGTCTTCCCGCAGGCTGAGATAGTGGGAATCCGGGACACNTGGGC 1707
Db 1410 ATCGAGAGTGTCTTCCCGCAGGCTGAGATAGTGGGAATCCGGGACACNTGGGC 1469
1708 CGGGCTGTAGATCAGGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAAACCTCTGA 1767
Db 1470 CGGGCTGTAGATCAGGTTTGCCCACTTGATTGAGAAAGATGTTCCAGTGAGGAAACCTCTGA 1529
1768 TCTNTCAGCCCCCAACACCCCAATTTGGCCCAACACTGNTTNGCCCTCGGGGTGTGAG 1827
Db 1530 TCTNTCAGCCCCCAACACCCCAATTTGGCCCAACACTGNTTNGCCCTCGGGGTGTGAG 1589
1828 AAATNTAGCGCAAGGCACTTTTAAACGTGGATTTGTTTAAAGAGCTCTCCAGGCCCCAC 1887
Db 1590 AAATNTAGCGCAAGGCACTTTTAAACGTGGATTTGTTTAAAGAGCTCTCCAGGCCCCAC 1649
1888 CAAGAGGTTGGATCACCTCAGTCGGAAGAAAAATTAATTTCTTTCAGGTTTAAAAA 1946
Db 1650 CAAGAGGTTGGATCACCTCAGTCGGAAGAAAAATTAATTTCTTTCAGGTTTAAAAA 1708

RESULT 5

US-09-261-855-1
; Sequence 1, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2224
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-261-855-1

Query Match

66.9%; Score 1301; DB 3; Length 2224;

; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 5
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-5

Query Match 85.8%; Score 1669.2; DB 3; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 268 CCAGGCAAGCCATCATCAAGCTGAATGCGCCACCAAGTTGGAGAACCATGCGCCTGAAGTCT 327
DB 30 CCGGGGAGCCATCATCAAGCTGAATGCGCCACCAAGTTGGAGAACCATGCGCCTGAAGTCT 89
QY 328 CTTACATCCCGATGAGCAGATAGCACAGGAGCCTGAGAAATGGCGCGGAGGGGGCTTTG 387
DB 90 CTTACATCCCGATGAGCAGATAGCACAGGAGCCTGAGAAATGGCGCGGAGGGGGCTTTG 149
QY 388 GCTCTCGGGTCAAGCCCGCCAGGCTCAGCTGTGGCAGCGGGGGCCCAAGCAAGCAGC 447
DB 150 GCTCTCGGGTCAAGCCCGCCAGGCTCAGCTGTGGCAGCGGGGGCCCAAGCAAGCAGC 209
QY 448 AGCAAGTGGACATCCCGCTCGCTCGCTGCTGCGCCACCAAGTATGGGTGCGCATATTG 507
DB 210 AGCAAGTGGACATCCCGCTCGCTGCTGCGCCACCAAGTATGGGTGCGCATATTG 269
QY 508 GCAAGGAGGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCCAAGATAGACGTG 567
DB 270 GCAAGGAGGGGGCCACCATCCGCAACATCAACAAACAGACCCAGTCCCAAGATAGACGTG 329
QY 568 ATAGGAGGAGACCGAGGTGAGCTGAAAGCCATCAAGTGTGACCTCAAGCCCTGAGG 627
DB 330 ATAGGAGGAGACCGAGGTGAGCTGAAAGCCATCAAGTGTGACCTCAAGCCCTGAGG 389
QY 628 GCTGCTCCTCCGCTTGAAGATGATCTTGAGATTATGATGAAGAGGCTTAAGGACCA 687
DB 390 GCTGCTCCTCCGCTTGAAGATGATCTTGAGATTATGATGAAGAGGCTTAAGGACCA 449
QY 688 AAACGGCTGACGAGGTTCCTGAGATCCTGCGCCCAATTAATCTTTAGGGCGTCTCA 747
DB 450 AAACGGCTGACGAGGTTCCTGAGATCCTGCGCCCAATTAATCTTTAGGGCGTCTCA 509
QY 748 TTGGCAAGGAGGACCGACCTGAGAGGTAGAGCAAGATACCGAGACAAATCAACA 807
DB 510 TTGGCAAGGAGGACCGACCTGAGAGGTAGAGCAAGATACCGAGACAAATCAACA 569
QY 808 TCTCTCTGTTCAAGACCTTACCTTTACCAACCTGAGAGGACCATCACTGTGAAGGGG 867
DB 570 TCTCTCTGTTCAAGACCTTACCTTTACCAACCTGAGAGGACCATCACTGTGAAGGGG 629
QY 868 CCATCGAATTTGTCAGGGCGGAGCAGGAATATGAAGAAAGTTCCGGAGGCGCTATG 927
DB 630 CCATCGAATTTGTCAGGGCGGAGCAGGAATATGAAGAAAGTTCCGGAGGCGCTATG 689
QY 928 AGAATGATGTGCTGATGCTCTCACTGATCCCTGCGCTGACCTGGCTGCTGTAG 987
DB 690 AGAATGATGTGCTGATGCTCTCACTGATCCCTGCGCTGACCTGGCTGCTGTAG 749
QY 988 GTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 1047
DB 750 GTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 809
QY 1048 CTCCTATAGTCTCTTTATGAGGTCCCGACAGAGATGCTGAGTGTATTATCCCGG 1107
DB 810 CTCCTATAGTCTCTTTATGAGGTCCCGACAGAGATGCTGAGTGTATTATCCCGG 869
QY 1108 CCCAGGAGTGGGCGCATCATCGCAAGAGGGGCGAGCAGATCAAGAGCTCTCCCGT 1167
DB 870 CCCAGGAGTGGGCGCATCATCGCAAGAGGGGCGAGCAGATCAAGAGCTCTCCCGT 929

QY 1168 TTGCCAGGCGCTCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTCTGTATGG 1227
DB 930 TTGCCAGGCGCTCATCAAGATTGCAACCCGAAACACCTGACTCCAAAGTTCTGTATGG 989
QY 1228 TTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAAATCTATGGCAAAAC 1287
DB 990 TTATCATCACTGGACCGCCAGAGGCCCAATTCAGGCTCAGGGAAGAAATCTATGGCAAAAC 1049
QY 1288 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG 1347
DB 1050 TCAAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG 1109
QY 1348 TGCAGCATCAGCAGCTGGCCGGTCAATTTGGCAAAAGTGGAAACCGGTGAACGAGTTGC 1407
DB 1110 TGCAGCATCAGCAGCTGGCCGGTCAATTTGGCAAAAGTGGAAACCGGTGAACGAGTTGC 1169
QY 1408 AGAATTTGACGCGAGCTGAGTGTAGTACCAAGAGACCCAGACCCCTGATGAGAACGACC 1467
DB 1170 AGAATTTGACGCGAGCTGAGTGTAGTACCAAGAGACCCAGACCCCTGATGAGAACGACC 1229
QY 1468 AGGTCACTCTGAAAATCATCGGACATTTCTATGCCAGTCAAGTGGTCAACGGAAGATCC 1527
DB 1230 AGGTCACTCTGAAAATCATCGGACATTTCTATGCCAGTCAAGTGGTCAACGGAAGATCC 1289
QY 1528 GAGACATCTGCCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGATTAACAGGCCCCAGG 1587
DB 1290 GAGACATCTGCCCCAGGTTAAGCAGCAGCATCAGAAGGGACAGATTAACAGGCCCCAGG 1349
QY 1588 CACGAGGAGGAGTACCAAGCCCTCCCTGTCCTTNGAGTCCAGGACAAACACGGGAGAA 1647
DB 1350 CACGAGGAGGAGTACCAAGCCCTCCCTGTCCTTNGAGTCCAGGACAAACACGGGAGAA 1409
QY 1648 ATCGAGGTGTCTCTCTCCCGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGC 1707
DB 1410 ATCGAGGTGTCTCTCTCCCGCAGGCTGAGAATGAGTGGGAATCCGGGACACNTGGGC 1469
QY 1708 CCGGCTGTAGATCAGGTTTGGCCACTTGAATTGAGAAAGATGTTCCAGTGAAGAACCTG 1767
DB 1470 CCGGCTGTAGATCAGGTTTGGCCACTTGAATTGAGAAAGATGTTCCAGTGAAGAACCTG 1529
QY 1768 TCTNTCAGCCCCCAACACCCACCTGAGTGGCCCAACCTGTTGAGTGGGCTGCTGAG 1827
DB 1530 TCTNTCAGCCCCCAACACCCACCTGAGTGGCCCAACCTGTTGAGTGGGCTGCTGAG 1589
QY 1828 AAATTNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGGCCCCAC 1887
DB 1590 AAATTNTAGCGCAAGGCACTTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGGCCCCAC 1649
QY 1888 CAAGAGGTGATCACACCTCAGTGGGAAGAAAAATAAAATTTCTTCAGGTTTTTAAA 1946
DB 1650 CAAGAGGTGATCACACCTCAGTGGGAAGAAAAATAAAATTTCTTCAGGTTTTTAAA 1708

RESULT 4

US-09-899-651-5
; Sequence 5, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709

1 GCTGTAGCGAGGGCTGGGGGCTGCTCTGCTCCCTTCTGTCGGGCTCGGCGCTCAGC 60
61 CCACCCAGAGGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGTTAGGAGCGGAAAA 120
61 CCACCCAGAGGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGTTAGGAGCGGAAAA 120
121 TTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTTGGACAGGCTGTGGCTC 180
121 TTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTACTTGGACAGGCTGTGGCTC 180
181 AGTATGTTAGTGTAGAGAACTGTAGAGAACTGTAGAGAACTGTAGAGAACTGTAGAG 240
181 AGTATGTTAGTGTAGAGAACTGTAGAGAACTGTAGAGAACTGTAGAGAACTGTAGAG 240
241 ATGTCACTATTCACCCGAGGAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
241 ATGTCACTATTCACCCGAGGAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
301 AGTTGGAGAAACCATGCTTGAAGTCTCTACATCCCGGATGAGCAGATAGCAGAGGAC 360
301 AGTTGGAGAAACCATGCTTGAAGTCTCTACATCCCGGATGAGCAGATAGCAGAGGAC 360
361 CTGAGAAATGGCGCGGAGGGGCTTTGGCTCTCGGGTTCAGCCCGGAGGGCTCACCTG 420
361 CTGAGAAATGGCGCGGAGGGGCTTTGGCTCTCGGGTTCAGCCCGGAGGGCTCACCTG 420
421 TGGCAGCGGGGCGCCAGCAGCAGCAAGTGGACATCCCGCTTCCGGCTCTCTGCTGTC 480
421 TGGCAGCGGGGCGCCAGCAGCAGCAAGTGGACATCCCGCTTCCGGCTCTCTGCTGTC 480
481 CCACCCAGTATGTGGTGCCATTTATGGCAAGGAGGGGCGCACCATCCGCAACATCAAA 540
481 CCACCCAGTATGTGGTGCCATTTATGGCAAGGAGGGGCGCACCATCCGCAACATCAAA 540
541 AACAGACCCAGTGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAGAGTGT 600
541 AACAGACCCAGTGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAGAGTGTAGAGTGT 600
601 CCATCAGTGTGACCTCCACCCCTGAGGGCTGCTCTCGCTTGTAAAGTGTCTTTGAGA 660
601 CCATCAGTGTGACCTCCACCCCTGAGGGCTGCTCTCGCTTGTAAAGTGTCTTTGAGA 660
661 TTATGCATAAAGAGGCTTAAGGACACAAACCGCTTGAAGGTTCCCTTGAAGTCTCTGG 720
661 TTATGCATAAAGAGGCTTAAGGACACAAACCGCTTGAAGGTTCCCTTGAAGTCTCTGG 720
721 CCCATAAATACTTTGTAGGGGCTCTCATTTGGCAAGGAGGAGCGAGGCTGAGAGAGT 780
721 CCCATAAATACTTTGTAGGGGCTCTCATTTGGCAAGGAGGAGCGAGGCTGAGAGAGT 780
781 AGCAGAGTACCGAGACAAATAATCACATCTCTCGTTGCAAGACCTTACCTTTTACAAC 840
781 AGCAGAGTACCGAGACAAATAATCACATCTCTCGTTGCAAGACCTTACCTTTTACAAC 840
841 CTGAGAGGACCATCACTGTGAAGGGGCGCATCGAGAAATTTGTGAGGGCGGAGGAGAA 900
841 CTGAGAGGACCATCACTGTGAAGGGGCGCATCGAGAAATTTGTGAGGGCGGAGGAGAA 900
901 TAATGAAGAAATTCGGGAGGCTATGAGAAATGATGTGCTGCCATGAGCTCTCACCTGA 960
901 TAATGAAGAAATTCGGGAGGCTATGAGAAATGATGTGCTGCCATGAGCTCTCACCTGA 960
961 TCCCTGGCTGACCTGGCTGTAGTGTCTTTTCCAGCTTCCATCCAGCGAGTCCCGC 1020
961 TCCCTGGCTGACCTGGCTGTAGTGTCTTTTCCAGCTTCCATCCAGCGAGTCCCGC 1020
1021 CGCTCCAGCAGCGTACTTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGC 1080
1021 CGCTCCAGCAGCGTACTTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCGAGC 1080
1081 AGAGATGTGTAGGTTTATCCCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
1081 AGAGATGTGTAGGTTTATCCCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140

1141 GGACGACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGACACCCG 1200
1141 GGACGACATCAACAGCTCTCCCGGTTTCCAGCGCTCCATCAAGATTGACACCCG 1200
1201 AACACCTGACTCCAAAGTTTCTGATGTTATCTACTGACCCAGAGGCGCAATTCA 1260
1201 AACACCTGACTCCAAAGTTTCTGATGTTATCTACTGACCCAGAGGCGCAATTCA 1260
1261 AGGCTCAGGAAAGAACTCTATGGCAACTCAAGGAGGAGAACTCTTTGGTCCCAGGAG 1320
1261 AGGCTCAGGAAAGAACTCTATGGCAACTCAAGGAGGAGAACTCTTTGGTCCCAGGAG 1320
1321 AAGTGAAGTGGAGACCCACATACGTCTGCCAGCATCAGCAGCTGGCGGGTCAATTGSCA 1380
1321 AAGTGAAGTGGAGACCCACATACGTCTGCCAGCATCAGCAGCTGGCGGGTCAATTGSCA 1380
1381 AAGGTGAAAGAAACCGTGAACAGGTTGAGAAATTTGACGGCAGCTGAGGTGTGTACCAA 1440
1381 AAGGTGAAAGAAACCGTGAACAGGTTGAGAAATTTGACGGCAGCTGAGGTGTGTACCAA 1440
1441 GAGACAGACCCCTGTATGAGAAACAGCAGGTTGAGAAATCATCGGACATTTCTATG 1500
1441 GAGACAGACCCCTGTATGAGAAACAGCAGGTTGAGAAATCATCGGACATTTCTATG 1500
1501 CCAGTCAAGTGGCTCAACGGAGAGATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATC 1560
1501 CCAGTCAAGTGGCTCAACGGAGAGATCCGAGACATCTCGGCCAGGTTAAGCAGCAGCATC 1560
1561 AGAAGGAGCAGAGTAAACAGGCGCCAGGCAACGAGAGAGTGAACAGGCGCTCTCTGCT 1620
1561 AGAAGGAGCAGAGTAAACAGGCGCCAGGCAACGAGAGAGTGAACAGGCGCTCTCTGCT 1620
1621 TNGAGTCCAGACAAACCGGCGAGAAATCGAGAGTGTCTCTCCCGCAGGCGCTGAGA 1680
1621 TNGAGTCCAGACAAACCGGCGAGAAATCGAGAGTGTCTCTCCCGCAGGCGCTGAGA 1680
1681 ATGAGTGGGAATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGGCCACTTGATTGA 1740
1681 ATGAGTGGGAATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGGCCACTTGATTGA 1740
1741 GAAAGATGTTCCAGTGAAGAACCCCTGATCTTNTCAGCCCCAAACACCCAAATTTGGGCC 1800
1741 GAAAGATGTTCCAGTGAAGAACCCCTGATCTTNTCAGCCCCAAACACCCAAATTTGGGCC 1800
1801 AACACTGNTGCCCTCGGGGTGTAGAAATTTAGCCAGGAGCTTTTAAACGTTGAT 1860
1801 AACACTGNTGCCCTCGGGGTGTAGAAATTTAGCCAGGAGCTTTTAAACGTTGAT 1860
1861 TGTTTAAAGAACTCTCCAGGCGCCACCCNAGAGGGTGGATCACCTCAGTGGGAGAAA 1920
1861 TGTTTAAAGAACTCTCCAGGCGCCACCCNAGAGGGTGGATCACCTCAGTGGGAGAAA 1920
1921 AATAAAATTTCTTCCAGGTTTAAAA 1946
1921 AATAAAATTTCTTCCAGGTTTAAAA 1946

RESULT 3
US-09-061-709-5
; Sequence 5, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538

Db 301 AGTTGGACACCATGTCCTGAGGTCCTCTACATCCCGATAGAGCAGATAGCAGGGAC 360
Qy 361 CTGAGAAATGGGGCGGAGGGGCTTTGGCTCTCGGGGTACGCCCGGCGAGGGCTCACCTG 420
Db 361 CTGAGAAATGGGGCGGAGGGGCTTTGGCTCTCGGGGTACGCCCGGCGAGGGCTCACCTG 420
Qy 421 TGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCTTCCTGGCTCTGGTGC 480
Db 421 TGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCTTCCTGGCTCTGGTGC 480
Qy 481 CCACCCAGTATGTGGGTGCCATTATTTGGCAAGAGGGGGCCACCATCCGCAACATCAAA 540
Db 481 CCACCCAGTATGTGGGTGCCATTATTTGGCAAGAGGGGGCCACCATCCGCAACATCAAA 540
Qy 541 AACACAGCCAGTCCAAGATAGACGTGCATAGGAGAGAGAACGAGGTGCAGCTGAAAAG 600
Db 541 AACACAGCCAGTCCAAGATAGACGTGCATAGGAGAGAGAACGAGGTGCAGCTGAAAAG 600
Qy 601 CCATCAGTGTGCACTCCACCCCTCAGGGCTGCTCCTCGCTTGTAAAGATGATCTTGAGA 660
Db 601 CCATCAGTGTGCACTCCACCCCTCAGGGCTGCTCCTCGCTTGTAAAGATGATCTTGAGA 660
Qy 661 TTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCTCTGAAGATCTGG 720
Db 661 TTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCTCTGAAGATCTGG 720
Qy 721 CCATATAACTTTGTAGGGGCTCTCATTTGGCAAGAGGAGAACGAGGTGCAGCTGAAAAG 780
Db 721 CCATATAACTTTGTAGGGGCTCTCATTTGGCAAGAGGAGAACGAGGTGCAGCTGAAAAG 780
Qy 781 AGCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCTTTTACAAAC 840
Db 781 AGCAAGATACCGAGACAAAATCACCATCTCCTCGTTGCAAGACCTTACCTTTTACAAAC 840
Qy 841 CTGAGAGGACCATCAGTGAAGGGGCCATCGAGATTTGTCAGGGCCGAGCAGGAAA 900
Db 841 CTGAGAGGACCATCAGTGAAGGGGCCATCGAGATTTGTCAGGGCCGAGCAGGAAA 900
Qy 901 TAATGAAGAAAGTTCCGGAGGCTATGAGATGATGCTGGCGCATGAGTCTCACCTGA 960
Db 901 TAATGAAGAAAGTTCCGGAGGCTATGAGATGATGCTGGCGCATGAGTCTCACCTGA 960
Qy 961 TCCTGGCTGAACCTGGCTGTAGTGTCTTTTCCAGCTTCATCCAGCGAGTCCCGC 1020
Db 961 TCCTGGCTGAACCTGGCTGTAGTGTCTTTTCCAGCTTCATCCAGCGAGTCCCGC 1020
Qy 1021 CGCTCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCTTTATGCAAGCTCCCGAGC 1080
Db 1021 CGCTCCAGCAGCTTACTGGGGCTGCTCCCTATAGCTCTTTATGCAAGCTCCCGAGC 1080
Qy 1081 AGGAGATGTCAGAGTGTATATCCCGCCAGCAGTGGGGCCATCATCGCAAGAGG 1140
Db 1081 AGGAGATGTCAGAGTGTATATCCCGCCAGCAGTGGGGCCATCATCGCAAGAGG 1140
Qy 1141 GGCAGACATCAAAAGCTCTCCGGTTTCCAGCGCTCCATCAAGATGCAACACCG 1200
Db 1141 GGCAGACATCAAAAGCTCTCCGGTTTCCAGCGCTCCATCAAGATGCAACACCG 1200
Qy 1201 ABAACCTGACTCCAAAGTTGATGGTTATCATCTGAGCGCGCAGAGGCCCAATCA 1260
Db 1201 ABAACCTGACTCCAAAGTTGATGGTTATCATCTGAGCGCGCAGAGGCCCAATCA 1260
Qy 1261 AGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAACTTCTTGTCCAGGAGG 1320
Db 1261 AGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAACTTCTTGTCCAGGAGG 1320
Qy 1321 AAGTGAAGCTGAGACCCACATAGCTGTGCCAGCATCAGCAGTGGCGGGTCATTGGCA 1380
Db 1321 AAGTGAAGCTGAGACCCACATAGCTGTGCCAGCATCAGCAGTGGCGGGTCATTGGCA 1380
Qy 1381 AAGTGGAAAAAGGCTGAACGATGTCAGAAATTTGACGCGCAGCTGAGTGGTAGTACCAA 1440

Db 1381 AAGTGGAAAAAAGCGTGAACGAGTTGCAGAAATTTTACGCGCAGCTGAGTGGTAGTACCAA 1440
Qy 1441 GAGACACAGACCCCTGATGAGAAACGACAGGTCATCGTGAATAATCATCGGACATTTCTATG 1500
Db 1441 GAGACACAGACCCCTGATGAGAAACGACAGGTCATCGTGAATAATCATCGGACATTTCTATG 1500
Qy 1501 CCAGTCAAGTGGCTCAACGGAAGATCCGAGACATCTCTGGCCAGGTTTAAAGCAGCAGATC 1560
Db 1501 CCAGTCAAGTGGCTCAACGGAAGATCCGAGACATCTCTGGCCAGGTTTAAAGCAGCAGATC 1560
Qy 1561 AGAAGGACAGAGTAACGAGGCCAGCAGCGAGGAAGTGAACGAGCAGGCTCCCT 1620
Db 1561 AGAAGGACAGAGTAACGAGGCCAGCAGCGAGGAAGTGAACGAGCAGGCTCCCT 1620
Qy 1621 TNGAGTCCAGGACAAACGAGGCGAGAAATCGAGAGTGTCTCTCCGCGGAGGCTCGAGA 1680
Db 1621 TNGAGTCCAGGACAAACGAGGCGAGAAATCGAGAGTGTCTCTCCGCGGAGGCTCGAGA 1680
Qy 1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA 1740
Db 1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA 1740
Qy 1741 GAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCCAAAACACCCCAATTTGGGCC 1800
Db 1741 GAAAGATGTTCCAGTGAAGAACCTGATCTNTCAGCCCCAAAACACCCCAATTTGGGCC 1800
Qy 1801 AACACTGTNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGCACCTTTTAAACGTGGAT 1860
Db 1801 AACACTGTNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGCACCTTTTAAACGTGGAT 1860
Qy 1861 TGTTTAAAGAGTCTCCAGGCCCCACCAAGAGGGTGGATCAACCTCAGTGGGAAGAAA 1920
Db 1861 TGTTTAAAGAGTCTCCAGGCCCCACCAAGAGGGTGGATCAACCTCAGTGGGAAGAAA 1920
Qy 1921 AATAAAATTTCTTCAGGTTTAAAA 1946
Db 1921 AATAAAATTTCTTCAGGTTTAAAA 1946

RESULT 2

US-09-899-651-7
; Sequence 7, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 538
; CURRENT APPLICATION NUMBER: US/09/899,651
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/061,709
; PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-7

Query Match 99.7%; Score 1941; DB 4; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCTGTAGCGGAGGGCTGGGGGCTGCTCTCTCCCTTCTTGGCGCTGGGCTCAGC 60
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 06:34:14 ; Search time 129.138 Seconds
(without alignments)
8362.646 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtagcggagggtctggg.....attcttcagggttttaaaa 1946

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1941	99.7	1946	3	US-09-061-709-7	Sequence 7, Appli
2	1941	99.7	1946	4	US-09-899-651-7	Sequence 7, Appli
3	1669.2	85.8	1708	3	US-09-061-709-5	Sequence 5, Appli
4	1669.2	85.8	1708	4	US-09-899-651-5	Sequence 5, Appli
5	1301	66.9	2224	3	US-09-261-855-1	Sequence 1, Appli
6	674.2	34.6	1740	4	US-09-643-597-347	Sequence 347, App
7	674.2	34.6	1740	4	US-09-542-615A-347	Sequence 347, App
8	674.2	34.6	1740	4	US-09-606-421B-347	Sequence 347, App
9	672.6	34.6	4159	3	US-09-061-709-4	Sequence 4, Appli
10	672.6	34.6	4159	4	US-09-899-651-4	Sequence 4, Appli
11	672.6	34.6	4181	4	US-09-643-597-175	Sequence 175, App
12	672.6	34.6	4181	4	US-09-480-884A-175	Sequence 175, App
13	672.6	34.6	4181	4	US-09-542-615A-175	Sequence 175, App
14	672.6	34.6	4181	4	US-09-606-421B-175	Sequence 175, App
15	585.4	30.1	3283	3	US-09-061-709-8	Sequence 8, Appli
16	585.4	30.1	3283	4	US-09-899-651-8	Sequence 8, Appli
17	570.6	29.3	3412	3	US-09-061-709-6	Sequence 6, Appli
18	570.6	29.3	3412	4	US-09-899-651-6	Sequence 6, Appli
19	53.6	2.8	7218	1	US-08-232-463-14	Sequence 14, Appli
20	45.6	2.3	598	4	US-09-669-751-39	Sequence 39, Appli
21	40.6	2.1	2850	1	US-08-503-172-5	Sequence 5, Appli
22	40.6	2.1	2850	2	US-09-135-211-5	Sequence 5, Appli
23	39.6	2.0	1982	4	US-09-016-434-1067	Sequence 1067, Ap
24	39.6	2.0	1982	4	US-09-825-497A-40	Sequence 40, Appl
25	39.6	2.0	1987	4	US-09-825-497A-39	Sequence 39, Appli
26	37.6	2.0	1987	4	US-09-517-779-1	Sequence 1, Appli
27	37.6	1.9	1896	4	US-09-343-011B-4	Sequence 4, Appli

28	1.9	2608	4	US-09-154-750A-75	Sequence 75, Appli
29	1.9	2608	4	US-09-665-479A-11	Sequence 11, Appli
30	1.8	555	4	US-09-252-991A-6251	Sequence 6251, Ap
31	1.8	1644	4	US-09-252-991A-6420	Sequence 6420, Ap
32	1.8	2088	4	US-09-252-991A-6439	Sequence 6439, Ap
33	1.8	2202	4	US-09-252-991A-6131	Sequence 6131, Ap
34	1.8	2463	4	US-09-252-991A-6775	Sequence 6775, Ap
35	1.8	9862	4	US-09-691-861A-3	Sequence 3, Appli
36	1.8	765	4	US-09-252-991A-16348	Sequence 16348, A
37	1.8	867	4	US-09-252-991A-15936	Sequence 15936, A
38	1.8	505	4	US-09-621-976-15639	Sequence 15639, A
39	1.8	1386	2	US-08-910-731-5	Sequence 5, Appli
40	1.8	2517	4	US-09-894-998A-38	Sequence 38, Appli
41	1.8	2655	4	US-09-963-137-139	Sequence 139, App
42	1.8	2655	4	US-09-963-137-183	Sequence 183, App
43	1.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
44	1.8	1926	4	US-09-249-585A-2	Sequence 2, Appli
45	1.8	1926	4	US-09-410-399-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-061-709-7
; Sequence 7, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 7
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-7

Query Match 99.7%; Score 1941; DB 3; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GCTGTAGCGAGGGGCTGGGGGGCTGCTCTGTCCCTTCCTTGGCGGCTGGCGGCTCAGC	60
DB	1	GCTGTAGCGAGGGGCTGGGGGGCTGCTCTGTCCCTTCCTTGGCGGCTGGCGGCTCAGC	60
QY	61	CCACCCAGAGCGGGGCTGGGGGGAGTGCTCAGCTTCCCGGTTAGAGCGGGA	120
DB	61	CCACCCAGAGCGGGGCTGGGGGGAGTGCTCAGCTTCCCGGTTAGAGCGGGA	120
QY	121	TTCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGTGGCTC	180
DB	121	TTCAATCCGAATATTTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTGTGGCTC	180
QY	181	AGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCAGAGTGAACCGGAGTGAACCGGAGTGTGA	240
DB	181	AGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACCCAGAGTGAACCGGAGTGTGA	240
QY	241	ATGTACCTATTCCACCGGAGCAGACGAGCAAGCATCATGAAGCTGAATGGCCACC	300
DB	241	ATGTACCTATTCCACCGGAGCAGACGAGCAAGCATCATGAAGCTGAATGGCCACC	300
QY	301	AGTTGGAGAACCATGCGCTTCTTACATCCCGGATGAGACATGACAGGAC	360

This Page Blank (uspto)

XX 28-JUN-2001; 2001WO-US021065.
XX 28-JUN-2000; 2000US-00606421.
XX 02-AUG-2000; 2000US-00630940.
XX 21-AUG-2000; 2000US-00643597.
XX 15-SEP-2000; 2000US-00662786.
XX 09-OCT-2000; 2000US-00685696.
XX 12-DEC-2000; 2000US-00735705.
XX 07-MAY-2001; 2001US-00850716.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
XX Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX WPI; 2002-090513/12.
XX P-PSDB; ABB75053.
XX Polynucleotides encoding lung tumor polypeptides, useful for treating
XX lung cancer or stimulating an immune response.
XX Claim 1; Page 367; 374pp; English.
XX The present invention describes human lung tumour proteins. Human lung
XX tumour proteins and polynucleotides have cytostatic and immunostimulant
XX activities, and can be used in vaccine production. Compositions
XX comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX proteins, T cell populations, or antigen presenting cells that express
XX the lung tumour proteins are useful for treating lung cancer or
XX stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX ABB75070 represent sequences used in the exemplification of the present
XX invention
XX Sequence 1743 BP; 527 A; 406 C; 418 G; 392 T; 0 U; 0 Other;
Query Match 34.6%; Score 674.2; DB 6; Length 1743;
Best Local Similarity 66.7%; Pred. No. 4.6e-177;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
QY 103 GGGTTAGGACGGGAAAATCAATCCGAATATTCACCCAGCTCCGATGGGAGTAC 162
DB 230 GCGAAAGGATTCGGAACCTTCAGATACGAATATTCGCCCTCATTTACAGTGGGAGTGC 289
QY 163 TGGACAGCCTGCTGCTCAGTATGTACAGTAGAAGCTGTGAGCAAGTGAACACCGAGA 222
DB 290 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGAGCTGTGAGCAAGTGAACATGACT 349
QY 223 GTGACAGCGCAGTGGTGAATGTCACTATTCCAAACGGGAGCAGACCGGCAAGCCATCA 282
DB 350 CGGAACCTGCAGTTGTAATGTAACTTATCCAGTAAAGGACCAAGCTAGACAAGCACTAG 409
QY 283 TGAAGCTGAATGGCCACAGTTGGAGAACCATATGCTCCAGCTCTCTACATCCCGATG 342
DB 410 ACAAACTGAATGGATTCAGTTAGAGAAATTCACCTTGAAGTAGCCTATATTCCTGATG 469
QY 343 ACAGATACCAAGGACCTGAGATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTTCAGC 402
DB 470 AAACGGCGCCCGAGCAAAACCCCTTTGACAGACGCCCGAGGTGCGCGGGGGCTTGGCGAGA 529
QY 403 CCCCGCAGGGCTCACTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 462
DB 530 GGGGCTCCTCAAGCAGGGGGTCTCCAGGATCCGATATCCAAAGCAGAAACCATGTGATTCG 589
QY 463 CCTTCGGCTCTGGTGGCCACCCAGTATGGTGGTGGCCATTTATTCAGAGGAGGGGGCCA 522
DB 590 CTCTGGGCTGCTGGTTCCGACCAATTTTGGAGGCCATCATAGGAAAGGAGGTGCCA 649
QY 523 CCATCCGACATCACAAACAGACCCAGTCCAAAGATAGACGTGCTATAGGAAAGGAGAAC 582
DB 650 CCATTCGGAAACATCACAAACAGACCCAGTCTAAATTCGATGTCCACCGTAAAGAAATG 709

QY 583 CAGGTGAGCTGAAAAAGCCATCAGTGTGCATCCACCCCTGAGGGCTGCTCCTCCGCTT 642
DB 710 CGGGGGCTGCTGAGAGTCGATTACTATCTCTCTACTCTCTGAGGCACTCTCTCGGCTT 769
QY 643 GTAAGATGATCTTTGGAGATTATCATAAAGAGCTAAGGACACCAAAACGGCTCACAGG 702
DB 770 GTAAGTCTATTCTGGAGATTATGCATAAGGAAGCTCAAGATATAAAATTCACAGAAGA 829
QY 703 TTCCCTGAAAGATCTTGGCCCATTAATTAACCTTTAGGGCGCTCTCATTTGGCAAGAGAC 762
DB 830 TCCCTTTGAAGATTTTAGCTCATATAAATCTTTGTGGACGCTCTATTGGTAAAGAGAA 889
QY 763 GGAACCTGAAAGAGGTAGAGCAAGATACCGAGACAAAATACCATCTCTCTGTTGCAAG 822
DB 890 GAATCTTTAAAAAATTTAGCAGACACACACACTAAATACAGATATCTCCTATGCAAG 949
QY 823 ACCTTACCTTTTACACCTCTGAGAGGACCATCACTGTGAAGGGGGCCCATCGAGAAATGTT 882
DB 950 AATTGACGCTGTATATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009
QY 883 GAGGGCCGAGCAGGAAATTAATGAGAAAGTTCCGGAGGCTATGAGAAATGATGTGGCTG 942
DB 1010 CCAAGCTGAGGAGGAGATCATGAAGAAAATCAGGGAGTCTTATGAAAATGATTAATGCTT 1069
QY 943 CCATGA-----GCTCTCACTGATCCCTGGCCCTGAACCTGGCTGCTGTAGGTCTTTTCC 996
DB 1070 CTATGAATCTCAAGACATTTAATCTGGATTAATCTGAACGCTTGGTCTGTGTCC 1129
QY 997 CAGCTTCATCCAGCGCAGTCCCGCGCTCCCGAGGGTTACTTGGGGCTGCTCCCTATA 1056
DB 1130 CACCACTTCAGGGATGCCACTCCACCTCAGGGCCCCCTTCAGCCATGATCTCTCCCT 1189
QY 1057 GCTCCTTTATGACGCTCCCGAGCAGGAGTGTGCGAGTGTATATCCCGCCCAAGGAG 1116
DB 1190 ACCCGCAGTTGAGCAATTCAGAAACGGAGATGTTCTCTGTTTATCCAGCTCTATCAG 1249
QY 1117 TGGGGCCATCATCGCAAGAGGGGAGCAGCATCAACACAGCTCTCCCGGTTTGCAGCG 1176
DB 1250 TCGGTGCCATCATCGCAGCAGGGCCAGCAGCATCAACAGCTTTCTCGCTTGTGGAG 1309
QY 1177 CTTCCATCAAGATTGCACCCCGAAACACCTGACTCCAAAGTTCGTATGGTTATCAUCA 1236
DB 1310 CTTCAATTAAGATTGCTCCAGCGAAGCACCAGATGCTTAAAGTGAGGATGGTATATCA 1369
QY 1237 CTGAGCCCGAGAGCCCAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGGAGG 1296
DB 1370 CTGACACCAAGAGGCTCAGTTCAAGGCTCAGGGGAAGATTTATGAAAAAATTAAGAG 1429
QY 1297 AGAATCTTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCAT 1356
DB 1430 AAACTTTGTTAGTCTTAAAGAGAGGTGAACCTTGAAGCTCATATCAGAGTGCCATCCT 1489
QY 1357 CAGCAGCTGGCGGGTCAATTTGGCAAGGTGGAAGAACCGGTGAACGAGTTGCAAGATTGA 1416
DB 1490 TTGCTGCTGGCAGAGTTATTTGAAAAGGAGGCAAAACCGTGAATGAATTCAGAAATTTG 1549
QY 1417 CGGCACTGAGGTGGTAGTACCAAGAGACCAAGCCCTGATGAGACGACCAAGGTCTATCG 1476
DB 1550 CAAAGTCAGAGAGTTTGTCTGTCGACACACCTGATGAGAAATGACCAAGTGTG 1609
QY 1477 TGAAATCATCGACATTTCTATGTCAGTTCAGATGGCTCAACCGAAGATCCGAGACATCC 1536
DB 1610 TCAAAATTAAGTGGTCACTTCTATGCTTGCAGGTTGCCAGAGTGGCCAGAGAAAATTCAGAAATTC 1669
QY 1537 TGGCCCGAGTTAAGCAGCAGCATCAGAAG 1565
DB 1670 TGACTCAGGTAAGCAGCAGCAACCAACAG 1698

Search completed: July 24, 2004, 07:28:34
Job time : 725.632 secs

CC	predetermined cut-off value, and determining the presence of a cancer in the patient. (I) and (II) are useful in pharmaceutical compositions, e.g. vaccines. (I) is useful as a marker to indicate the presence or absence of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to ABP61992 represent sequences used in the exemplification of the present invention	CC
XX	Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;	XX
SY	Query Match 34.6%; Score 674.2; DB 6; Length 1740;	SY
CC	Best Local Similarity 66.7%; Pred. No. 4.6e-177;	CC
CC	Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;	CC
QY	103 GGGTTAGGACCGGAAAAATTCAAATCCGAAATATTCACCCAGCTCCGATGGGAAGTAC 162	QY
DB	230 GGCAAAGGATTCCGAAATTCAGATACGAAATATCCCGCTCATTTACAGTGGGAGGTGC 289	DB
QY	163 TGGACACCTGCTGGCTCAGTATGCTACAGTACGAACTGTGAGCAAGTGAACACCGAGA 222	QY
DB	290 TGGATATGTTACTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGTACT 349	DB
QY	223 GTGAGACGGCAGTGGTCAATGTCACTATTTCCAAACCGGAGCAGACAGGCAAGCCATCA 282	QY
DB	350 CGGAACTGCAGTTGTAAATGTAACTTATCCAGTAAGGACCAAGCTAGACAAGCACTAG 409	DB
QY	283 TGAAGCTGAATGGCCACAGCTTGGAGAA CCAATGCCCTGAAGCTCTCTACATCCCGATG 342	QY
DB	410 ACAAACTGAATGATTTTCAGTTAGAGAAATTCACCTTGAAGTAGGCTTATATCCCTGATG 469	DB
QY	343 AGCAGATAGCACAGGGACCTGAGAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTCAGC 402	QY
DB	470 AAACGGCGCCACGAAAAA CCCCCTTCAGCAGACCCCGAGSTCGCCGGGGCTTGGSCAGA 529	DB
QY	403 CCGCCACGGGCTACCTGTGGCAGCGGGGCCCCAGCAGAGCAGAGTGGACATCC 462	QY
DB	530 GGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGTTATCCAAAGCAGAAACCATGTGATTTGC 589	DB
QY	463 CCTTTGGCTCTGGTCCCAACCCAGTATCTGGGTGCCATTTATGGCAGAGGGGGCCA 522	QY
DB	590 CTCTGGCGCTGCTGGTTCCCA CCCCATTGTTGGAGCCATCATAGGAAAGAGGTGCCA 649	DB
QY	523 CCATCCGCAACATCACAAAA CAGACCCAGTCCAAAGATAGACGTGCATAGGAAGGAGAACG 582	QY
DB	650 CCATTCCGAA CATCACCAACAGACCCAGCTCTAAATCGATGTCCACCGTAAAGAAAAATG 709	DB
QY	583 CAGTGCAGCTGAAAAAGCCATCAGTGTGACTCCACCCCTGAGGGCTGCTCCTCCGCTT 642	QY
DB	710 CGGGGGCTGCTGAGAAGTCGATTACTATCTCTACTCTCTGAGGACCTCTGCGGCTT 769	DB
QY	643 GTAGATCATCTTCGAGATTATGCATAAAGAGGCTTAAGGACACCAAAACCGCTGACGAGG 702	QY
DB	770 GTAAGCTATTCTGGAGATTATGATATAGGAGAGCTCAAGATATATAAATTCACAGAAAGAG 829	DB
QY	703 TTCCCTGAAGATCCTGCCCATATAATCTTTGTAGGGGCTCTCATTTGGCAAGAGGAC 762	QY
DB	830 TCCCTCTCAAGATTTTAGCTCATATAACTTTGTTGGACGCTCTTAITGGTAAAGAAAGGAA 889	DB
QY	763 GGAACCTGAAGAAGGTAGACGAATACCGAGACAAAATCACCATCTCCTCGTTGCAAG 822	QY
DB	890 GAAATCTTAAAAAAATTTAGCAAGACACAGACATATAAATCACGATATCTCCATTGCAAG 949	DB
QY	823 ACCTTACCTTTTACAACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATGTT 882	QY
DB	950 AATTGACGCTGTATTAATCCAGAACGCACTATTACAGTTAAAGGCAATGTTGAGACATGTG 1009	DB
QY	883 GCAGGCGCGAGCAGGAAATATATGAAGAAGTTCCGGAGGCCCTATGAGAAATGATGCTGCTG 942	QY
DB	1010 CCAAAAGCTGAGGAGGAGATCATGAAGAAATTCAGGAGCTTTATGAAAATGATATTGCTT 1069	DB
QY	943 CCATGA-----GCTCTCACCTGTACCTTGGCCTGAACCTGGCTGCTGTAGGCTCTTTTCC 996	QY
DB	1070 CTATGAATCTTCAAGCACATTTAAATCTCTGGATTAAATCTGAACGCTTTGGTCTGTCTCC 1129	DB

```
Db 290 TGGTAGTATTACTAGTCCAGTATCGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACT 349
Qy 223 GTGAGAGCGGAGTGGTGAATGTCACTTATCCAAACCGGAGCAGACCAAGGCAAGCCATCA 282
Db 350 CGGAAACTGCGAGTGTGAATGTAACTTATCCAGTAAGGACCAAGCTAGACAAGCACTAG 409
Qy 283 TGAAGCTCAATGGGCCACAGTGTGAGAACCAATGCCCTGAAAGTCTCTCTACATCCCCATG 342
Db 410 ACAAACCTGAATGATTTTCAGTTAGAGATTTTCACCTTGAAGTAGCCCTATATCCCTGATG 469
Qy 343 AGCAGATAGCACAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGC 402
Db 470 AAACGGCGCGCCAGCAAAACCCCTTGACAGACCCCGAGGTGCGCGGGGCTTGGGAGAG 529
Qy 403 CCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCAAGCAGCAGCAAGGTGACATCC 462
Db 530 GGGGCTCCTCAAGCAGGGGTCTCCAGGATCCGATATCCAAAGCAAAACCATGTGATTTGC 589
Qy 463 CCCTTCGGCTCCTGGTGCCCAACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCA 522
Db 590 CTCTGGCCCTGCTGGTTCCCAOCCAAATTTGTTGGAGCCATCATAGNAAGAAAGGTGCCA 649
Qy 523 CCAATCCGACATCACAAACAGACCCAGTCCAGATAGACGTGTCATAGGAAGGAGAACG 582
Db 650 CCATTCGGAAACATCACCAACAGACCCAGTCTAAATTCGATGCCACCGTAAAGAAATG 709
Qy 583 CAGGTGCGAGCTGAAAAAGCCATCAGTGTGCACCTCCACCCCTGAGGGTGTCTCTCGCTT 642
Db 710 CGGGGCTGCTGAGAGTCAATTAATCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 769
Qy 643 GTAAGATGATCTTTGGAGATTTATGATAAAGAGGCTAAGGACACCAAAACGGCTGACGAG 702
Db 770 GTAAGTCTATTCTGGAGATTATGCAATAGGAAGCTCAAGATATAAATTCACAGAAGAGA 829
Qy 703 TTCCCTGAAGATCTCGGCCATTAATAAATTTGTAGGGGTCTCATTTGGCAAGGAAGGAC 762
Db 830 TCCCTTGAAGATTTTGTAGTCTAATAAATCTTTGTGAGCTCTTATTGGTAAGAAGGAA 889
Qy 763 GGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAATACCATCTCTCTGTTGCAAG 822
Db 890 GAAATCTTAATAAATTTGAGCAAGACACAGACACTAAATACAGATATCTCAATTCAGG 949
Qy 823 ACCTTACCTTTACAACTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTG 882
Db 950 AATTGACGTGTATATCCAGAACGCACTATTACAGTTAAGCAATGTTGAGACATGTG 1009
Qy 883 GCAGGCGCAGCAGGAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAATGATGGCTG 942
Db 1010 CCAAGCTGAGGAGGAGATCATGAAGAAATCAGGGAGTCTTATGAAAATGATATTGCTT 1069
Qy 943 CCATGA-----GCTCTCACCTGATCCCTGGCTGAACTGGCTGTGTAGTCTTTTCC 996
Db 1070 CTATGAATCTTCAAGCACATTAATTTCTCGGATTAATCTGAACGCTTGGGTCTGTTC 1129
Qy 997 CAGTCTATCCAGCGAGTCCGCGCCCTCCAGCAGCGTTACTGGGGCTGCTCCCTATA 1056
Db 1130 CACCCACTTCAGGATGCCACCTCCACCTCAGGGCCCCCTTCAGCCATGACTCTCCCT 1189
Qy 1057 GCTCTTTATGAGGCTCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGCCAGGAG 1116
Db 1190 ACCCGAGTTTGAGCAATCAGAAAACGAGACTGTTTCATCTGTTTATCCAGCTCATCAG 1249
Qy 1117 TGGGGCCCATCATCGCAAGAGGGGCGAGCATCAAAACAGTCTCCCGGTTTGCAGG 1176
Db 1250 TCGGTGCCATCATCGCAAGAGGGGCCAGCATCAAGCAGTTTCTCGCTTGTGGAG 1309
Qy 1177 CTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTTAAGTGAGGATGTTGATATCA 1236
Db 1310 CTTCAATTAAGATTGCTCCAGCGGAAGCACCAGATGCTTAAGTGAGGATGTTGATATCA 1369
Qy 1237 CTGGACCCCGAGGCCCAATTCAGGCTCAGGGAGGAATCTATGGCAAACTCAAGGAGG 1296
```

```
Db 1370 CTGACCAACACAGAGGCTCAGTTCAAGGCTCAGGAGAGATTTATGGAATAATTAAGAAG 1429
Qy 1297 AGAATCTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTTGTCAGCAT 1356
Db 1430 AAAATCTTGTAGTCTTAAGAGAGAGTGAACCTTGAAGCTCATATCAGAGTGCCATCCT 1489
Qy 1357 CAGCAGCTGCGCGGCTCAITTTGGCAAAAGGTGGAATAAAACGGTGAACGAGTTTGCAGAAATTGA 1416
Db 1490 TTGCTGTGTCAGAGTTATTGGAAAAGGAGGCAAAACGGTGAATGAATTTTCAAGATTTGT 1549
Qy 1417 CGCAGCTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGACGACAGGTCATCG 1476
Db 1550 CAAGTCGAGAAGTTGTTGTCCTCGTGACCAAGACACTGATGAGAATGACCAAGTGGTTG 1609
Qy 1477 TGAATAATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCC 1536
Db 1610 TCAAAATAATCTGCTCACTTCTATGCTTGCCAGGTTGCCAGAGAAAATTCAGGAATTC 1669
Qy 1537 TGSCCCAGGTTAAGCAGCAGCATCAGAAG 1565
Db 1670 TGACTCAGGTAAGACGACCAACAACAG 1698

RESULT 13
ABQ92440
ID ABQ92440 standard; cDNA; 1740 BP.
XX
AC ABQ92440;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human lung cancer associated cDNA sequence SEQ ID NO:347.
XX
KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine; gene;
XX ss.
XX Homo sapiens.
XX
PN WO200247534-A2.
XX
PD 20-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-US047576.
XX
PR 12-DEC-2000; 2000US-00735705.
XX
PR 07-MAY-2001; 2001US-00850716.
XX
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR WPI; 2002-583465/62.
XX
DR P-PSDB; ABP61917.
XX
PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
PS Claim 1; Page 337; 381pp; English.
XX
CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
```


CC methods. The present sequence was used as the representative sequence
CC from a human clone which was used in homology searches to identify the
CC clone. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in CD-ROM format directly from
CC EPO
XX
SQ Sequence 833 BP; 223 A; 225 C; 215 G; 166 T; 0 U; 4 Other;
Query Match 37.7%; Score 732.8; DB 4; Length 833;
Best Local Similarity 96.9%; Pred. No. 1.5e-193;
Matches 802; Conservative 0; Mismatches 14; Indels 12; Gaps 5;
QY 537 ACATAACAGACCCAGTCCAGATAGACGTGCATAGGAGGAGAACGCGAGTGCAGCTGAA 596
Db 1 ACATAACAGACCCAGTCCAGATAGACGTGCATAGGAGGAGAACGCGAGTGCAGCTGAA 60
QY 597 AAAGCCATCAGTGTGACTCCACCCCTGAGGCTGCTCCTCCGCTTGAAGATGATCTTG 656
Db 61 AAAGCCATCAGTGTGACTCCACCCCTGAGGCTGCTCCTCCGCTTGAAGATGATCTTG 120
QY 657 GAGATTATGCATAAAGAGGCTAAGGACACCAAAACGCTGACGAGTTCCCTGAGATC 716
Db 121 GAGATTATGCATAAAGAGGCTAAGGACACCAAAACGCTGACGAGTTCCCTGAGATC 180
QY 717 CTGGCCCATTAATTAACCTTTGAGGCGTCTCATTTGGCAAGGAGGACGGAACTGAAGAAG 776
Db 181 CTGGCCCATTAATTAACCTTTGAGGCGTCTCATTTGGCAAGGAGGACGGAACTGAAGAAG 240
QY 777 GTAGAGCAGATACCGGACGACAAATAACCATCTCTCGTTGCAAGACCTTACCTTTAC 836
Db 241 GTAGAGCAGATACCGGACGACAAATAACCATCTCTCGTTGCAAGACCTTACCTTTAC 300
QY 837 AACCTTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGCGAGGGCGGAGCAG 896
Db 301 AACCTTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGCGAGGGCGGAGCAG 360
QY 897 GAAATTAATGAAGAAATTCGGAGGCGCTATGAGAAATGATGTGCTGCCATGAGC----- 950
Db 361 GAAATTAATGAAGAAATTCGGAGGCGCTATGAGAAATGATGTGCTGCCATGAGCCTGCAG 420
QY 951 TCTACCTCATCCTCGGCTGACCTGACCTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGC 1010
Db 421 TCTACCTCATCCTCGGCTGACCTGACCTGGCTGCTGTAGTCTTTTCCAGCTTCATCCAGC 480
QY 1011 GCAGTCCCGCGCTCCCGACGAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGAG 1070
Db 481 GCAGTCCCGCGCTCCCGACGAGCGTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCGAG 540
QY 1071 GCTCCCGAGCAGAGATGTGAGTGTGATTTATCCCGCCGAGGCGAGTGGCGCCCATCATC 1130
Db 541 GCTCCCGAGCAGAGATGTGAGTGTGATTTATCCCGCCGAGGCGAGTGGCGCCCATCATC 600
QY 1131 GGCAAGAGGGGCGAGCAGATCAACAGCTCTCCCGTTTGGCAGCGCTTCATCAAGATT 1190
Db 601 GGCAAGAGGGGCGAGCAGATCAACAGCTCTCCCGTTTGGCAGCGCTTCATCAAGATT 660
QY 1191 GCACACCGGAACACCTGACTCCAAAGTTTCGTATGTTATCATCACTGGA-CCGCCAGA 1249
Db 661 GCACACCGGAACACCTGACTCCAAAGTTTCGTATGTTATCATCACTGGAAGCCGCCAGA 720
QY 1250 GGCCCAATTCAGGCTCAGGGAAGAAATCATGGCAAACTC-AAGGAGGAGAACTT--CTT 1306
Db 721 NGCCCAATTCAGGCTCANGGAAGAAATCATGGCAAACTC-AAGGAGGAGAACTTCTTT 780
QY 1307 TGGTCCCAAGGAGGAAGT--GAGGCTGGAGACCCACATACGTGTGCCA 1352
Db 781 GGTCCCAAGGAGGAAGTGAAGCTTTGGAGACCCACATACCTGTGCCA 828
RESULT 11
AAC66035
ID AAC66035 standard; cDNA; 1740 BP.
XX

AAC66035;
21-FEB-2001 (first entry)
Human lung cancer-associated cDNA antigen L523S.
Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
vaccine; detection; ss.
Homo sapiens.
W0200061612-A2.
19-OCT-2000.
03-APR-2000; 2000WO-US008896.
02-APR-1999; 99US-00285479.
17-DEC-1999; 99US-00466396.
30-DEC-1999; 99US-00476496.
10-JAN-2000; 2000US-00480884.
22-FEB-2000; 2000US-00510376.
(CORI-) CORIXA CORP.
Wang T, Fan L;
WPI; 2000-628399/60.
P-PSDB; AAB11365.
Isolated polypeptide comprising an immunogenic portion of a lung tumor
protein is used for detecting and monitoring progression of lung cancer
in a patient.
Claim 1a; Page 258-259; 261pp; English.
This invention describes a novel isolated polypeptide (I) which
comprising an immunogenic portion of a lung tumor protein or variant (P2)
which have cytostatic activity. The polypeptides and polynucleotides are
used in compositions and vaccines to inhibit the development of cancer,
especially lung cancer, in a patient. Methods described in the invention
can be used to monitor the progression of a cancer by carrying out the
detection at subsequent time points and comparing the results from the
different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
are treated with P2, polynucleotides encoding P2 or antigen presenting
cells expressing P2 and then administered to the patient to inhibit
development of cancer
Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;
Query Match 34.6%; Score 674.2; DB 3; Length 1740;
Best Local Similarity 66.7%; Pred. No. 4.6e-177;
Matches 980; Conservative 0; Mismatches 483; Indels 6; Gaps 1;
QY 103 GGGTTAGAGCGCGGAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAGTAC 162
Db 230 GCGAAAGATTGCGAACTTCAGATACGAAATATCCCGCTCAATTACAGTGGGAGTGC 289
QY 163 TGGACAGCGCTGCTGCTCAGTATGGTACAGTAGAGAACTGTGAGCACTGACACCGAGA 222
Db 290 TGGATAGTTTACTAGTCCAGTATGGAGTGGTGGAGCTGTGAGCACTGACACTGACT 349
QY 223 GTGAGACGGCAGTGGTGAATCTCACTTATTCACACCGGAGCAGCAGGCAAGCATCA 282
Db 350 CGGAACTGCAGTTGTAATGTAACCTTATTCAGTAGGACCAAGCTAGACAGCACTAG 409
QY 283 TGAAGCTGAATGGCCACAGTTGGAGAACCATGCGCTGAAGGTCTCTATATCCCGGATG 342
Db 410 ACAACTGAATGGATTTTCAGTTAGAGAAATTTACCTTGAAGTAGCTTATATCCCTGATG 469
QY 343 AGCAGATAGCACAGGACCTGAGAAATGGCGCCGAGGGGGCTTTGGCTCTCGGGGTGAGC 402
Db 470 AAACGGCGCCAGCAAAACCCCTTGACAGAGCCCGAGGTGCGCGGGGGCTTGGCGAGA 529

27-SEP-2000; 2000US-0235834P.
 29-SEP-2000; 2000US-0236327P.
 29-SEP-2000; 2000US-0236367P.
 29-SEP-2000; 2000US-0236368P.
 29-SEP-2000; 2000US-0236369P.
 29-SEP-2000; 2000US-0236370P.
 02-OCT-2000; 2000US-0236802P.
 02-OCT-2000; 2000US-0237037P.
 02-OCT-2000; 2000US-0237038P.
 02-OCT-2000; 2000US-0237039P.
 13-OCT-2000; 2000US-0237040P.
 20-OCT-2000; 2000US-0239935P.
 20-OCT-2000; 2000US-0240960P.
 20-OCT-2000; 2000US-0241785P.
 20-OCT-2000; 2000US-0241809P.
 01-NOV-2000; 2000US-0244617P.
 17-NOV-2000; 2000US-0249299P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 XX
 (ROSE/) ROSEN C A.
 (RUBE/) RUBEN S M.
 (BARA/) BARASH S C.
 PA
 PA
 PA
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 WPI; 2001-147444/14.
 DR P-PSDB; ABUS5647.
 DR
 XX
 New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS
 Claim 1; SEQ ID NO 745; 402pp; English.
 XX
 CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABX731173-ABX74167 represent
 CC human novel polynucleotides of the invention
 XX
 SQ Sequence 1129 BP; 318 A; 296 C; 302 G; 210 T; 0 U; 3 Other;

Query Match 55.0%; Score 1069.8; DB 7; Length 1129;
 Best Local Similarity 98.7%; Pred. No. 2e-287;
 Matches 1084; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 855 ACTGTGAAGGGGCCATCGAGAAATTTGCGGGCCGAGCGAGGAATAATGAAGAAAGTT 914
 DB 1 ACTGTGAAGGGGGCCATCGAGAAATTTGCGGGCCGAGCGAGGAATAATGAAGAAAGTT 60
 QY 915 CGGAGGGCCTATGAGATGATGTGGCTGCCATGAGC-----TCTCAGCTGATCCCTGGC 968
 DB 61 CGGAGGGCCTATGAGATGATGTGGCTGCCATGAGCCTGCGAGCTTCACCTGATCCCTGGC 120
 QY 969 CTGAACCTGGCTGCTGTAGGTCCTTTTCCAGCTTCATCCAGCGAGTCCCGCCGCTCCC 1028
 DB 121 CTGAACCTGGCTGCTGTAGGTCCTTTTCCAGCTTCATCCAGCGAGTCCCGCCGCTCCC 180
 QY 1029 AGCAGCGTTACTGGGCTGCTCCCTTATAGCTTCCTTTATGCAAGGCTCCGAGCAGAGATG 1088

Db 181 AGCAGCGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAGATG 240
 QY 1089 GTGCAGGTGTTTATCCCGCCAGGAGTGCGGCCCATCATCGCAGAGAGGGGCGAGCAC 1148
 DB 241 GTGCAGGTGTTTATCCCGCCAGGAGTGCGGCCCATCATCGCAGAGAGGGGCGAGCAC 300
 QY 1149 ATCAAAAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCAACCCGAAACACCT 1208
 DB 301 ATCAAAAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCAACCCGAAACACCT 360
 QY 1209 GACTCCAAAGTTTGTATGTTTATCATCTGAGCCGCGAGGCCCAATTAAGGCTCAG 1268
 DB 361 GACTCCAAAGTTTGTATGTTTATCATCTGAGCCGCGAGGCCCAATTAAGGCTCAG 420
 QY 1269 GGAAGAATCTATGCGAAACTCAAGGAGGAGAACTTTCTTTGGTCCCGAGGAGGAGTAAG 1328
 DB 421 GGAAGAATCTATGCGAAACTCAAGGAGGAGAACTTTCTTTGGTCCCGAGGAGGAGTAAG 480
 QY 1329 CTGGAGACCCACATACGTGTGCGAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGA 1388
 DB 481 CTGGAGACCCACATACGTGTGCGAGCATCAGCAGCTGGCGGGTCAATTGGCAAGGTGGA 540
 QY 1389 AAAACGGTGAACGAGTTGCGAATTTGACGGGAGCTGAGTGGTAGTACCAAGAGACCGAG 1448
 DB 541 AAAACGGTGAACGAGTTGCGAATTTGACGGGAGCTGAGTGGTAGTACCAAGAGACCGAG 600
 QY 1449 ACCCTGTATGAGAACGACGAGGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAG 1508
 DB 601 ACCCTGTATGAGAACGACGAGGTTCATCGTGAATAATCATCGGACATTTCTATGCCAGTCAG 660
 QY 1509 ATGGGTCAACGAAAGATCCGAGACATCTTGGCCCGAGTTTAAGCAGCAGCATCAGAAAGGA 1568
 DB 661 ATGGGTCAACGAAAGATCCGAGACATCTTGGCCCGAGTTTAAGCAGCAGCATCAGAAAGGA 720
 QY 1569 CAGAGTTACCGCCCGCCAGCAGGAGGAGTGAACGAGCCCTCCCTGTCCCTTNGAGTCC 1628
 DB 721 CAGAGTTACCGCCCGCCAGCAGGAGGAGTGAACGAGCCCTCCCTGTCCCTTNGAGTCC 780
 QY 1629 AGGACAAACAAACGGGCGAGAAATCGAGAGTGTGCTCTCCCGGCGAGGCTGAGAAATGAGTGG 1688
 DB 781 AGGACAAACAAACGGGCGAGAAATCGAGAGTGTGCTCTCCCGGCGAGGCTGAGAAATGAGTGG 840
 QY 1689 GAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGCACATTTGATTGAGAAAGATG 1748
 DB 841 GAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGCACATTTGATTGAGAAAGATG 900
 QY 1749 TTCCAGTCAGGAACCTCATCTNTCAGCCCAACCAACCCACCAATTTGCCCAACACTGT 1808
 DB 901 TTCCAGTCAGGAACCTCATCTNTCAGCCCAACCAACCCACCAATTTGCCCAACACTGT 960
 QY 1809 NTGCCCTCCGGGGTGTGAGAAATTTAGCGCAAGGCATTTTAAACGTTGGATTGTTTAAA 1868
 DB 961 NTGCCCTCCGGGGTGTGAGAAATTTAGCGCAAGGCATTTTAAACGTTGGATTGTTTAAA 1020
 QY 1869 GAAGCTCTCCAGGCGCCACCAAGAGGGTGGATCACTCAGTGGGAGAAAAATAAAT 1928
 DB 1021 GAAGCTCTCCAGGCGCCACCAAGAGGGTGGATCACTCAGTGGGAGAAAAATAAAT 1080
 QY 1929 TTCCTTCAGGTTTAAAA 1946
 DB 1081 TTCCTTCAGGTTTAAAA 1098

RESULT 9
 AAK91969
 ID AAK91969 standard; cDNA; 833 BP.
 XX
 AC AAK91969;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human cDNA 5'-end sequence, SEQ ID NO: 429.
 XX

CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 55.0%; Score 1069.8; DB 4; Length 1129;
Best Local Similarity 98.7%; Pred. No. 2e-287; 5; Indels 6; Gaps 1;
Matches 1084; Conservative 3; Mismatches

QY 855 ACTGTCAAGGGGCCATCGAATATTGTTCAGGGCCGAGCAGGAAATATGAAGAAGTT 914
DB 1 ACTGTCAAGGGGCCATCGAATATTGTTCAGGGCCGAGCAGGAAATATGAAGAAGTT 60
QY 915 CGGAGGCGCTATGAGAAATGATGTGGCTGCATGAGC-----TCTACCTGATCCCTGGC 968
DB 61 CGGAGGCGCTATGAGAAATGATGTGGCTGCATGAGCCTGCGAGTCTCACCTGATCCCTGGC 120
QY 969 CTGAACCTGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCC 1028
DB 121 CTGAACCTGCTGCTGTAGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCC 180
QY 1029 AGCAGGCTTACTCGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGATG 1088
DB 181 AGCAGGCTTACTCGGGCTGCTCCCTATAGTCTCTTTATGAGGCTCCCGAGCAGATG 240
QY 1089 GTGAGGCTTTATCTCCCGCCAGGAGTGGGCGCTCATTCGCGAAGAGGGCAGCAC 1148
DB 241 GTGAGGCTTTATCTCCCGCCAGGAGTGGGCGCTCATTCGCGAAGAGGGCAGCAC 300
QY 1149 ATCAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGTTGACACCCGAAACACT 1208
DB 301 ATCAACAGCTCTCCGGTTTCCAGCGCTCCATCAAGTTGACACCCGAAACACT 360
QY 1209 GACTCCAAAGTTCTGATGTTATCATCACTGAGCGCGCAGAGGCCCAATCAAGCTCAG 1268
DB 361 GACTCCAAAGTTCTGATGTTATCATCACTGAGCGCGCAGAGGCCCAATCAAGCTCAG 420
QY 1269 GGAAGAATCTATGGCAACTCAAGAGGAGAACTCTTTTGTGCCAAGGAGGAGTGAAG 1328
DB 421 GGAAGAATCTATGGCAACTCAAGAGGAGAACTCTTTTGTGCCAAGGAGGAGTGAAG 480
QY 1329 CTGAGAGCCACATACGTTGTCAGCATCAGAGCTGGCGGCTCATTTGGCAAGAGTGA 1388
DB 481 CTGAGAGCCACATACGTTGTCAGCATCAGAGCTGGCGGCTCATTTGGCAAGAGTGA 540
QY 1389 AAAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGCCAG 1448
DB 541 AAAACGGTGAACGAGTTCAGAAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGCCAG 600
QY 1449 ACCCTGTATGAAACGACAGCTCATCTGTAATAATCATCGGACATTTCTATGCCAGTCA 1508
DB 601 ACCCTGTATGAAACGACAGCTCATCTGTAATAATCATCGGACATTTCTATGCCAGTCA 660
QY 1509 ATGGCTCAACGAAAGATTCGAGACATCTTGGCCCGCAGTTAAGCAGCAGCATCAGAGGGA 1568
DB 661 ATGGCTCAACGAAAGATTCGAGACATCTTGGCCCGCAGTTAAGCAGCAGCATCAGAGGGA 720
QY 1569 CAGAGTTAACCGGCCAGCAGCGAGGAAGTGAACAGCCCTCCCTGTCTCCCTTNGAGTCC 1628
DB 721 CAGAGTTAACCGGCCAGCAGCGAGGAAGTGAACAGCCCTCCCTGTCTCCCTTNGAGTCC 780
QY 1629 AGGACAAACCGGCGCAAAATCGAGTGTGCTCTCCCGGCGAGGCTGAGATGATGG 1688
DB 781 AGGACAAACCGGCGCAAAATCGAGTGTGCTCTCCCGGCGAGGCTGAGATGATGG 840
QY 1689 GAATCCGGGACANTGGCGGCTGTAGATGAGTTTGGCCACTTGTGATGAGAAAGATG 1748
DB 841 GAATCCGGGACANTGGCGGCTGTAGATGAGTTTGGCCACTTGTGATGAGAAAGATG 900

QY 1749 TTCCAGTGAGGAACCTGTATCTTTCAGCCGCCAAACACCCCAATGGCCCAACTGT 1808
DB 901 TTCCAGTGAGGAACCTGTATCTTTCAGCCGCCAAACACCCCAATGGCCCAACTGT 960
QY 1809 NTGCCCCCTCGGGTGTGAGAAATTTAGCGCAAGGCACTTTTAAAGCTGGATTGTTAAA 1868
DB 961 CTGCCCCCTCGGGTGTGAGAAATTTAGCGCAAGGCACTTTTAAAGCTGGATTGTTAAA 1020
QY 1869 GAAGCTTCTCAGGCCCCCAGAGAGGCTGGATCACACCTCAGTGGGAAGAAAAATAAAT 1928
DB 1021 GAAGCTTCTCAGGCCCCCAGAGAGGCTGGATCACACCTCAGTGGGAAGAAAAATAAAT 1080
QY 1929 TTCTTTTTCAGGTTTAAAA 1946
DB 1081 TTCTTTTTCAGGTTTAAAA 1098
RESULT 8
ABX73907
ID ABX73907 standard; DNA; 1129 BP.
XX
AC ABX73907;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polynucleotide #735.
XX
KW Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
EN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214888P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.

polynucleotides. The polypeptides and polynucleotides are useful in gene therapy for treating, inhibiting or preventing neural disorders, immune system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis), muscular disorders, respiratory diseases (e.g. nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders, gastrointestinal disorders, pulmonary disorders, cardiovascular disorders (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left heart syndrome), renal disorders (e.g. acute kidney failure and end-stage renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and leukaemia), inflammatory diseases (e.g. septic shock, bursitis and appendicitis), allergic reactions and conditions (e.g. asthma), blood related disorders (e.g. thrombosis, atherosclerosis and myocardial infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent human novel polynucleotides of the invention.

1192 BP: 343 A: 310 C: 314 G: 215 T: 0 U: 0 Other:

Query Match	55.8%	Score 1085.4;	DB 7;	Length 1182;
Best Local Similarity	98.9%	Pred. No. 9.1e-292;		
Matches 1102; Conservative	0;	Mismatches 6;	Indels 6;	Gaps 1;

839	CCCTGAGGAGGACCATCACTGTGAAGGGGCCCATCGAGAAATTGTTCCAGGGCCGACAGGA	898
18	CGCTGAGAGGACCATCACTGTGAAGGGGGCCCATCGAGAAATTGTTCCAGGGCCGACAGGA	77
899	AATAATGAAGAAAGTTTCGGGAGGCCCTATGAGAATGATGTGGCTGCCATGAGC-----TC	952
78	AATAATGAAGAAAGTTTCGGGAGGCCCTATGAGAATGATGTGGCTGCCATGAGCCTCAGTTC	137
953	TCACCTGATCCCTGGCCTGAAACCTGGCTGCTGTAGTCTTTTCCAGTCTTCATCCAGCGC	1012
138	TCACCTGATCCCTGGCCTGAAACCTGGCTGCTGTAGTCTTTTCCAGTCTTCATCCAGCGC	197
1013	AGTCCGCGCGCTCCCCAGCAGCGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGACGC	1072
198	AGTCCGCGCGCTCCCCAGCAGCGTTACTGGGGCTGCTCCCTATAGTCTCTTTATGACGC	257
1073	TCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGGCCCAGGCAGTGGGGCCCATCATCGG	1132
258	TCCCGAGCAGGAGATGGTGCAGGTGTTTATCCCGGCCCAGGCAGTGGGGCCCATCATCGG	317
1133	CAAGAAGGGGCGAGCATCAAAACAGCTCTCCCGGTTTGCAGCGCTCCCATCAAGATTGC	1192
318	CAAGAAGGGGCGAGCATCAAAACAGCTCTCCCGGTTTGCAGCGCTCCCATCAAGATTGC	377
1193	ACACCCGGAACACCTGACTTCAAAAGTTCGTATGGTTTATCATCACTGGAGCCGCCAGAGGC	1252
378	ACCACCCGAACACCTGACTTCAAAAGTTCGTATGGTTTATCATCACTGGAGCCGCCAGAGGC	437
1253	CCAAATTCAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCC	1312
438	CCAAATTCAGGCTCAGGGAAGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCC	497
1313	CAAGGAGGAAGTGAAGCTGAGAGCCCATACAGTGTGCCAGCATCAGCAGCTGSCCGGGT	1372
498	CAAGGAGGAAGTGAAGCTGAGAGCCCATACAGTGTGCCAGCATCAGCAGCTGSCCGGGT	557
1373	CATTGGCAAGGTTGGAAAACCGTGAACGAGTTTCAGAAATTTGACGGCAGCTGAGGTGGT	1432
558	CATTGGCAAGGTTGGAAAACCGTGAACGAGTTTCAGAAATTTGACGGCAGCTGAGGTGGT	617
1433	AGTACCAGAGAGCCAGACCCCTGATGAGACGACGAGGTCACTCGTAAATCATCGGACA	1492
618	AGTACCAGAGAGCCAGACCCCTGATGAGACGACGAGGTCACTCGTAAATCATCGGACA	677
1493	TTTCTATGCCAGTCCAGATGGCTCAACGGAAAGATCCGAGACATCTGCGCCAGGTTAAGCA	1552
678	TTTCTATGCCAGTCCAGATGGCTCAACGGAAAGATCCGAGACATCTGCGCCAGGTTAAGCA	737
1553	GCAGCATCAGAGGACGACAGTAAACCGGGCCGAGCACCGAGGAAGTGACACGCCCTCC	1612
738	GCAGCATCAGAGGACGACAGTAAACCGGGCCGAGCACCGAGGAAGTGACACGCCCTCC	797

Db 378 ACCACCGAAGACCTGACTCCAAAGTTCGATGGTATCATCACTGGACCGCAGAGGC 437
 QY 1253 CCAATTCAGGCTCAGGGAAGAAATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCC 1312
 Db 438 CCAATTCAGGCTCAGGGAAGAAATCTATGCAAACTCAAGGAGGAGAACTTCTTTGGTCC 497
 QY 1313 CAAGGAGGAAGTGAAGCTGGAGACCCACATACCTGTGCGAGCATCAGCAGCTGGCGGGT 1372
 Db 498 CAAGGAGGAAGTGAAGCTGGAGACCCACATACCTGTGCGAGCATCAGCAGCTGGCGGGT 557
 QY 1373 CATTCGCAAGGTGGAAAAACGGTGAACCGATGTCAGAAATTCACGCGCAGCTCAGGTGCT 1432
 Db 558 CATTCGCAAGGTGGAAAAACGGTGAACCGATGTCAGAAATTCACGCGCAGCTCAGGTGCT 617
 QY 1433 AGTACCAAGAGACCAAGACCCCTGATGAGAACGACAGGTCATCGTGAATAATCATCGGACA 1492
 Db 618 AGTACCAAGAGACCAAGACCCCTGATGAGAACGACAGGTCATCGTGAATAATCATCGGACA 677
 QY 1493 TTTCTATGCGAGTCAGATGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCA 1552
 Db 678 TTTCTATGCGAGTCAGATGCTCAACGGAAGATCCGAGACATCCTGGCCAGGTTAAGCA 737
 QY 1553 GCAGCATCAGAAGGGACAGAGTAACAGGCCAGGACGAGGAAGTGACACGCGCCCTCC 1612
 Db 738 GCAGCATCAGAAGGGACAGAGTAACAGGCCAGGACGAGGAAGTGACACGCGCCCTCC 797
 QY 1613 CTGTCCTTNGAGTCAGAGAACAAACGCGGAGAAATCGAGAGTGCTCTCCCGGCAG 1672
 Db 798 CTGTCCTTNGAGTCAGAGAACAAACGCGGAGAAATCGAGAGTGCTCTCCCGGCAG 857
 QY 1673 GCCTGAGATGATGGGAATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGGCCAC 1732
 Db 858 GCCTGAGATGATGGGAATCCGGGACACNTGGCGGCTGTAGATCAGGTTTGGCCAC 917
 QY 1733 TTGATGTGAAGATGTTTCCAGTAGGAAACCTGATCTNTAGCCCCAAACACCCACCA 1792
 Db 918 TTGATGTGAAGATGTTTCCAGTAGGAAACCTGATCTNTAGCCCCAAACACCCACCA 977
 QY 1793 ATTGGCCCAACTGTTGCGGCTGGGCTGTCAGAAATNTAGCGAAGGACATTTAA 1852
 Db 978 ATTGGCCCAACTGTTGCGGCTGGGCTGTCAGAAATNTAGCGAAGGACATTTAA 1037
 QY 1853 ACGTGGATGTTTAAAGAGCTCTCAGGCCCCACCAAGAGGCTGATCAGACCTCAGTG 1912
 Db 1038 ACGTGGATGTTTAAAGAGCTCTCAGGCCCCACCAAGAGGCTGATCAGACCTCAGTG 1097
 QY 1913 GGAAGAAAAATAAATTTCTTCAGGTTTAAAA 1946
 Db 1098 GGAAGAAAAATAAATTTCTTCAGGTTTAAAA 1131

RESULT 6

ID ABX73489 standard; DNA; 1182 BP.

AC AC ABX73489;

DT 18-MAR-2003 (first entry)

DE Human novel polynucleotide #317.

XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;

KW muscular disorder; respiratory disease; reproductive disorder;

KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;

KW hyperproliferative disorder; inflammatory disease; allergic reaction;

KW blood related disorder; cancer; immunosuppressive; antiinflammatory;

KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;

KW haemostatic; antiarteriosclerotic.

OS Homo sapiens.

XX US2002132753-A1.

FN

XX PD 19-SEP-2002.
 XX PF 17-JAN-2001; 2001US-00764864.
 XX PF 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0228688P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-147444/14.

P-PSDB; ABUS5229.

New polypeptides and nucleic acids, useful in gene therapy for treating, inhibiting or preventing e.g. neural, immune system, muscular, respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or renal disorders.

Claim 1; SEQ ID NO 327; 402pp; English.

The invention relates to human novel polypeptides and their associated

Db 1385 GCATCATTCGCAAGAGGGCCAGCACATCAACACATCTCCCTTTCCGCGAGCGCTCC 1444
QY 1182 ATCAAGATTGACACACCGGAAACCTGACTCCAAAAGTTGGTATGATCATCATCTGGA 1241
Db 1445 ATCAAGATTGCTCCACGAGAAACCTGACTCCAAAAGTTGGAATGGTTCATCATCTGGA 1504
QY 1242 CCGCAGAGGCCCAATTCAGGCTCAGGAGAGATCTATGCAAACTCAAGGAGGAAC 1301
Db 1505 CCCCAGAGGCTGAGTTCAAGGCTCAGGAGAGATTTATGGCAACTAAAAGAGAGAT 1564
QY 1302 TTCTTTGGTCCCAAGGAGAGTGAAGCTGAGACCCACACATACGTGTGCCAGCATCAGCA 1361
Db 1565 TTCTTTGGTCCCAAGGAGAGTGAAGCTGAGACCCACACATACGGTTCCGGCTCAGCA 1624
QY 1362 GCTGGCCGGGTCTATGSCAAAGTGGAAAACGGTGAACCGATTGACAGCA 1421
Db 1625 GCCGGCCGGTCTATCGGCAAGGCGCAAAACGGTGAATGAGCTGCAGAACTTGACTGCA 1684
QY 1422 GCTGAGGTGCTAGTACCAAGAGACCAAGACCCCTGATGAGAACCAAGCTCATCTGGA 1481
Db 1685 GCTGAGGTGCTAGTACCAAGAGACCAAGACCCCGATGAGAACCAAGCTCATCTGTAAG 1744
QY 1482 ATCATCGACATTTCTATGCCATGATGAGTGGCTCAACGGGAAGATCCGAGACATCTGGCC 1541
Db 1745 ATCATCGACATTTCTATGCCATGATGAGTGGCTCAACGGGAAGATCCGAGACATCTGGCT 1804
QY 1542 CAGTTTAAGCAGCAGCATCAGAGGACAGAGTAACAGGCCAGGACCGAGGAAGTGA 1601
Db 1805 CAAAGTTAAGCAGCAGCATCAGAGGACAGAGTAACAGGCCAGGACCGAGGAAGTGA 1864
QY 1602 -CCAGCCCTCCCTGTCCTTNGAGTCCAGGACCAACAGGGCAGAA----- 1647
Db 1865 CCCCCTGCTCTGCTCCATTTGGCTCCAGATCAGCAGGAGAACACAACTCGAGGG 1924
QY 1648 -----ATCGAGAGTGTCTCTCCCGGAGGCTGAGATGAGTGGGAATCCGGGAC 1699
Db 1925 GCGGGTGGAGGGCCGGTGTGTTTTCCAGCAGGCTGAGATGAGTGGGAATCAGGG-C 1983
QY 1700 ACNTGGCCGGCTGTAGATCAGGTTTGGCCACTTGAATGAGAAAGATTTCCAGTGAGG 1759
Db 1984 ATTTGGCCCTGTGAGATCAGGTTTGCACACTGTATGAGAACAAATTTCCAGTGAGG 2043
QY 1760 AACCTGTATCTNTACGCCCCCAACACCCCAATTTGGCCCAACTGNTGCCCTCGG 1819
Db 2044 AATCTGTATCTCTGCCCCCAA--TTGAGCCAGCTGGCCACAGCCACCCCTTGGAAAT 2101
QY 1820 GGTGTCAAAATNTAGCGCAAGGCATTTTAAACGTGATTTTAAAGAGCTCTCCA 1879
Db 2102 CACCAATTCATATGAGTGGTGTGTTTAAACGTGATTTGCT--TGAAGTTCTCCA 2159
QY 1880 GGCCCCCAAGAGGGTGCATCACACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGT 1939
Db 2160 GCCTCCATGAGGATGGTGCATCCAGTGGGGAAGAGAAATAAAATTTCTTCAGGT 2219
QY 1940 TTTA 1943
Db 2220 TTTA 2223

RESULT 5

AAS26148
ID AAS26148 standard; cDNA; 1182 BP.
XX
AC AAS26148;
XX
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 327.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;

secreted protein; rheumatoid arthritis; hyperproliferative disorder;
cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
cerebral ischaemia; angiogenesis; nervous system disorder;
Alzheimer's disease; infection; ocular disorder; corneal infection;
wound healing; epithelial cell proliferation; skin ageing; food additive;
preservative; antiproliferative.
Homo sapiens.

WO200155322-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US0001341.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-0189874P.

17-MAR-2000; 2000US-0190076P.

18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209467P.

28-JUN-2000; 2000US-0214886P.

30-JUN-2000; 2000US-0215135P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-0216800P.

11-JUL-2000; 2000US-0217487P.

11-JUL-2000; 2000US-0217496P.

14-JUL-2000; 2000US-0218290P.

26-JUL-2000; 2000US-0220963P.

14-AUG-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225213P.

14-AUG-2000; 2000US-0225214P.

14-AUG-2000; 2000US-0225266P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225447P.

14-AUG-2000; 2000US-0225757P.

14-AUG-2000; 2000US-0225758P.

18-AUG-2000; 2000US-0226279P.

22-AUG-2000; 2000US-0226681P.

22-AUG-2000; 2000US-0226868P.

22-AUG-2000; 2000US-0227182P.

30-AUG-2000; 2000US-0227009P.

01-SEP-2000; 2000US-0228924P.

01-SEP-2000; 2000US-0229287P.

01-SEP-2000; 2000US-0229343P.

01-SEP-2000; 2000US-0229344P.

05-SEP-2000; 2000US-0229345P.

05-SEP-2000; 2000US-0229509P.

06-SEP-2000; 2000US-0229513P.

06-SEP-2000; 2000US-0230437P.

08-SEP-2000; 2000US-0230438P.

08-SEP-2000; 2000US-0231242P.

08-SEP-2000; 2000US-0231243P.

08-SEP-2000; 2000US-0231244P.

08-SEP-2000; 2000US-0231413P.

08-SEP-2000; 2000US-0231414P.

08-SEP-2000; 2000US-0232080P.

12-SEP-2000; 2000US-0232081P.

14-SEP-2000; 2000US-0231968P.

14-SEP-2000; 2000US-0232397P.

14-SEP-2000; 2000US-0232398P.

14-SEP-2000; 2000US-0232399P.

14-SEP-2000; 2000US-0232400P.

14-SEP-2000; 2000US-0232401P.

14-SEP-2000; 2000US-0233063P.

Mon Jul 26 09:17:14 2004

Db 1081 CCCTGTCCTTCCAGTCCAGGACAAACACGGGACAGAAATCGAGAGTGTGCTCTCCCGGC 1140
 Qy 1671 AGGCTGAGATGAGTGGGATCCGGACACNTGGGCGGCTGTAGATCAGGTTTGCC 1730
 Db 1141 AGGCTGAGATGAGTGGGATCCGGACACCTGGGCGGCTGTAGATCAGGTTTGCC 1200
 Qy 1731 ACTTGATTGAGAAAGATGTTCCAGTGAGGACCCCTGATCTNCTCAGCCCCAAACACCCACC 1790
 Db 1201 ACTTGATTGAGAAAGATGTTCCAGTGAGGACCCCTGATCTCAGCCCCAAACACCCACC 1260
 Qy 1791 CAATTGGCCCCAACACTGTTNGCCCTCGGGGTGTGAGAAATTTAGCGCAAGCACTTTT 1850
 Db 1261 CAATTGGCCCCAACACTGTTNGCCCTCGGGGTGTGAGAAATTTAGCGCAAGCACTTTT 1320
 Qy 1851 AAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATCACACCTCAG 1910
 Db 1321 AAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATCACACCTCAG 1380
 Qy 1911 TGGGAGAGAAATAAAATTTCTTTCAGGTTTAAAA 1946
 Db 1381 TGGGAGAGAAATAAAATTTCTTTCAGGTTTAAAA 1416

RESULT 4

AAZ10617
 ID AAZ10617 standard; cDNA; 2224 BP.

XX AAZ10617;

DT 17-NOV-1999 (first entry)

XX cDNA encoding a murine c-myc coding region determinant binding protein.

XX c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
 KW endonucleolytic attack; half-life; breast cancer; colon cancer;
 KW pancreatic cancer; ss.

XX Mus musculus.

XX Key Location/Qualifiers
 FH 131..1864
 CDS /*tag= a

XX WO9946594-A2.

XX 16-SEP-1999.

XX 05-MAR-1999; 99WO-US004897.

XX 09-MAR-1998; 98US-0077372P.

XX (WISC) WISCONSIN ALUMNI RES FOUND.

XX Ross J;

XX WPI; 1999-551506/46.

XX P-FSDB; AAY30849.

XX Diagnosing presence or absence of a tumor in a human by examining c-myc
 PT coding region determinant-binding protein.

XX Example; Fig 1A-D; 79pp; English.

XX The present sequence encodes a murine c-myc coding region determinant
 CC binding protein (CRD-BP). The presence or absence of a tumor can be
 CC determined by determining the levels of CRD-BP present in the suspect
 CC tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack
 CC and so prolongs its half-life. The methods are used for diagnosing
 CC presence or absence of a tumor in a human, especially breast, colon and
 CC pancreatic cancer. They are also used to inhibit cancer cell growth

XX Sequence 2224 BP; 580 A; 609 C; 618 G; 417 T; 0 U; 0 Other;

SQ

Query Match 66.7%; Score 1297.8; DB 2; Length 2224;
 Best Local Similarity 83.4%; Pred. No. 0;
 Matches 1554; Conservative 0; Mismatches 277; Indels 33; Gaps 6;
 Qy 108 AGGAGCCGGAATTCAAATCCGAAATATTCACCCAGCTCCGATGGAAGTACTGGAC 167
 Db 365 AGGAGTCCGAAATACAGATCCGCAATATTCACCTCAGCTCCGATGGAAGTACTGAT 424
 Qy 168 AGCCTGCTGGCTCAGTATGTTAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 227
 Db 425 AGCCTGCTGGCTCAGTACGTTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 484
 Qy 228 AGCGCAGTGTGCAATGTCACTTATTCACACCGGGAGCAGACAGGCAAGCCATCATGAG 287
 Db 485 ACAGCGTGTGTGCACTGCTACCTTCTAACCAGGAGCAGACAGGCAAGCTATCATGAG 544
 Qy 288 CTGAATGGCCACCAAGTTGGAGAACATGCTTCAAGTCTTCTACATACCTGATGAGCAG 347
 Db 545 CTAAATGGCCATCAACTGGAGAACCATGCTTGAAGTCTTCTACATACCTGATGAGCAG 604
 Qy 348 ATAGCACAGGACCTCAGATGGCGCGCCGAGAGGGGCTTTGGCTCTCGGGGTACGCCCCG 407
 Db 605 ATAAACAAGGTCCTGAGAAATGGCGCTGCTGGAGGCTTTGGGTCTCGGGCCAGCCCCG 664
 Qy 408 CAGCGCTCACCTGTGCGAGCGGGGCCCCAGCCAGCAGCAGCAGTGGACATCCCCCTT 467
 Db 665 CAAGGGTCGCCCGTGGCAGCAGGGGCTCCAGCCAAAGCAGCAGCAGTGGACATCCCCCTC 724
 Qy 468 CGGCTCCTGTGCCCCACCGAGTATGTTGGTGCATTTATTCGCAAGGAGGGGCCACCATC 527
 Db 725 CGGCTCCTGTGCTTACGCGATGTATGAGGCGCTATCAATGGCAAGGAGGGTGCACCATC 784
 Qy 528 CGCAACATCAAAACAGACCCAGTCCAAAGTACAGCTGATGATGATGATGATGATGATGAT 587
 Db 785 CGAAACATCAAAACAGACCCAGTCCAAAGTACAGCTGATGATGATGATGATGATGATGAT 844
 Qy 588 GCAGCTGAAAAGCCATCAGTGTGCTCCACCCCTGAGGGGTGCTCTCGCTTGTGAAG 647
 Db 845 GCTCGGAGAGGCCATCAGCGTGCATTCACCCCTGAAAGCTGTCTCTCGCGGTGCAAG 904
 Qy 648 ATGATCTGGAGATTATGATTAAGAGGCTTAAGACACCAAAACGGCTGACGAGTGTCCC 707
 Db 905 ATGATCTGGAGATTATGATTAAGAGGCTTAAGACACCAAAACGGCTGACGAGTGTCCC 964
 Qy 708 CTGAGATCTCGGCCCAATAAATTTGTAGGGGCTCTCATTTGGCAAGAGAGGACCGAAC 767
 Db 965 CTGAAGATCTCGGCTCATPAACACTTCTCGGGCAGCTCATTTGGCAAGAGGCGGGAAC 1024
 Qy 768 CTGAAGAGGTAGACCAAGATACCGAGACAAATTCACATCTCTCTGTTGCAAGACCTT 827
 Db 1025 CTGAAGAGGTAGACCAAGATACCGAGACAAATTCACATCTCTCTGTTGCAAGACCTT 1084
 Qy 828 ACCCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCCCATCGAATTTGTTGACGG 887
 Db 1085 AGCTCTATAACCTTGAGAGGACCATCACTGTGAAGGGGGCCCATTTGAGAACTGTTGACGG 1144
 Qy 888 GCGGAGCGGAATTAATGAAGAAAGTTGCGGAGGCTCATGAGAATGATGCTGCTGCTGCTG 947
 Db 1145 GCGGAGCGGAGATCATGAAGAAAGTTGAGAGGCTTACGAGAACGACGCTGCGCGCATG 1204
 Qy 948 AGC-----TCATCAGCTGATCCCTGGCTGAAACCTGGCTGTGCTGCTGCTGCTGCTGCT 1001
 Db 1205 AGCTTGCACTCCCACTCATCTCTGGCTTAACTGGCTGTGAGGTCTCTTCTCCAGCT 1264
 Qy 1002 TCATCAGCGAGTCCCGCGGCTCCAGCAGGTTTACTGGGCTCTCTCCCTATAGCTCC 1061
 Db 1265 TCATCAGCGGCTGTCCCTCTCTCCAGCAGTGTCTCTGGGCTGCTCTCCCTATAGCTCC 1324
 Qy 1062 TTATTCAGGCTCCCGCAGCAGAGATGTTGAGGTTTATCCCGCCAGGAGGAGTGGGC 1121
 Db 1325 TTATTCAGGCTCCCGCAGCAGAGATGTTGAGGTTTATCCCGCCAGGAGGAGTGGGC 1384
 Qy 1122 GCCATCATCGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTGCCAGGCTCC 1181

Db 1470 CGGCTGTAGATCAGGTTTGCACCTTGAUTGAGAAAGATGTTCCAGTGAGGACCTGA 1529
QY 1768 TCTNTCAGCCCCAAACACCCCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTCAG 1827
Db 1530 TCINTCAGCCCCAAACACCCCAATTTGGCCCAACACTGNTGCCCTCGGGGTGTCAG 1589
QY 1828 AAATNTAGCGCAAGGCACCTTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGGCCCCAC 1887
Db 1590 AAATNTAGCGCAAGGCACCTTTTAAACGTGGATTGTTTAAAGAAAGCTCTCCAGGCCCCAC 1649
QY 1888 CAAGAGGTCGATCACACCTCAGTCAGTGGGAAGAAATAAATTTCCCTTCAGGTTTAAAA 1946
Db 1650 CAAGAGGTCGATCACACCTCAGTCAGTGGGAAGAAATAAATTTCCCTTCAGGTTTAAAA 1708

RESULT 3

AAK94782

ID AAK94782 standard; cDNA; 2780 BP.

XX AAK94782;

XX 06-NOV-2001 (first entry)

DE Human full-length cDNA, SEQ ID NO: 3886.

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EP1130094-A2.

XX PD 05-SEP-2001.

XX PF 07-JUL-2000; 2000EP-00114089.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000JP-00118774.

XX PR 02-MAY-2000; 2000JP-00183765.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR WPI; 2001-524255/58.

XX DR P-PSDB; AAM93826.

XX 830 Primers useful for synthesizing full length cDNA clones and their use

XX in genetic manipulation.

XX Claim 8; SEQ ID NO 3886; 1380pp + Sequence Listing; English.

XX The invention relates to primers for synthesising full length cDNA
XX clones. 830 cDNA molecules encoding a human protein have been isolated
XX and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX been determined. Primers for synthesising the full length cDNA are useful
XX for clarifying the function of the protein encoded by the cDNA. The full
XX length clones were obtained by construction of full length enriched cDNA
XX libraries that were synthesised by the oligo-capping method. The primers
XX enable the production of the full length cDNA easily without any special
XX methods. The present sequence is a full length human cDNA of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in CD-ROM format directly
XX from EPO

XX SQ Sequence 2780 BP; 768 A; 681 C; 671 G; 660 T; 0 U; 0 Other;

Query Match

Best Local Similarity 71.4%; Score 1389; DB 4; Length 2780;

Matches 1405; Conservative 99.2%; Pred. No. 0;

Mismatches 5; Indels 6; Gaps 1;

QY 537 ACAAAACAGACCCCAAGATAGAGCTCATAGGAAGGAGAACGAGGTGCAGCTGAA 596

Db 1 ACACAAACAGACCCCAAGATAGAGCTCATAGGAAGGAGAACGAGGTGCAGCTGAA 60
QY 597 AAAGCCATCAGTGTGACATCCACCCCTGAGGCTGCTCCTCCGCTGTGAAGATGATCTTG 656
Db 61 AAAGCCATCAGTGTGACATCCACCCCTGAGGCTGCTCCTCCGCTGTGAAGATGATCTTG 120
QY 657 GAGATTATGATTAAGAGGCTTAAGGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 716
Db 121 GAGATTATGATTAAGAGGCTTAAGGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 180
QY 717 CTGGCCCAATAAATCTTTGTAGGGCTCTCATNTGGCAAGAGGACCGAACCTGGAAG 776
Db 181 CTGGCCCAATAAATCTTTGTAGGGCTCTCATNTGGCAAGAGGACCGAACCTGGAAG 240
QY 777 GTAGACCAAGATACCGAGACAAATAACCATCTCCTCGTTGCAAGACCTTACCTTTAC 836
Db 241 GTAGACCAAGATACCGAGACAAATAACCATCTCCTCGTTGCAAGACCTTACCTTTAC 300
QY 837 AACCTTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATGTTGCAAGGCCGAGCAG 896
Db 301 AACCTTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATGTTGCAAGGCCGAGCAG 360
QY 897 GAAATAATGAAGAAAGTTGGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCCTGCAG 950
Db 361 GAAATAATGAAGAAAGTTGGGAGGCTATGAGAAATGATGTGGCTGCCATGAGCCTGCAG 420
QY 951 TCTCACTGATCCCTGGCTGAACCTGCTGCTGAGGCTGCTCCTATAGCTCCTTTATGCGAG 1010
Db 421 TCTCACTGATCCCTGGCTGAACCTGCTGCTGAGGCTGCTCCTATAGCTCCTTTATGCGAG 480
QY 1011 GCAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTTATGCGAG 1070
Db 481 GCAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTGCTCCTATAGCTCCTTTATGCGAG 540
QY 1071 GCTCCCGAGCAGGATGCTGAGGCTGTTTATCCCGCCGAGGCTGAGTGGGCCATCATC 1130
Db 541 GCTCCCGAGCAGGATGCTGAGGCTGTTTATCCCGCCGAGGCTGAGTGGGCCATCATC 600
QY 1131 GGCAGAGAGGGGCGAGCAGATCAAAACAGCTCTCCCGTTTCCAGCGCTCCATCAGATT 1190
Db 601 GGCAGAGAGGGGCGAGCAGATCAAAACAGCTCTCCCGTTTCCAGCGCTCCATCAGATT 660
QY 1191 GCACACCCGAAACACCTGACTCCAAAGTTGCTATGTTTATCATCACTGACCGCCAG 1250
Db 661 GCACACCCGAAACACCTGACTCCAAAGTTGCTATGTTTATCATCACTGACCGCCAG 720
QY 1251 GCCCAATTCAGGCTCAGGAGGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 1310
Db 721 GCCCAATTCAGGCTCAGGAGGATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGT 780
QY 1311 CCCAAGAGGAGGAGTGAAGCTGGAGACCCACATAGCTGTGCCAGCATCAGAGCTGCGCG 1370
Db 781 CCCAAGAGGAGGAGTGAAGCTGGAGACCCACATAGCTGTGCCAGCATCAGAGCTGCGCG 840
QY 1371 GTCAATGGCAAGGTTGAAACCGTGAACCGAGTTGAGAAATTTGACGGCAGCTGAGGTG 1430
Db 841 GTCAATGGCAAGGTTGAAACCGTGAACCGAGTTGAGAAATTTGACGGCAGCTGAGGTG 900
QY 1431 GTAGTACCAAGAGCAGACCCCTGATGAGAACGACGACGAGTCACTGTAATAATCATCGGA 1490
Db 901 GTAGTACCAAGAGCAGACCCCTGATGAGAACGACGAGTCACTGTAATAATCATCGGA 960
QY 1491 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCCCCAGGTTAG 1550
Db 961 CATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCCGAGACATCTGCCCCAGGTTAG 1020
QY 1551 CAGCAGCATCAGAGGAGCAGAGTAAACAGGCCCGGACGAGGAGTGAACGAGCCCT 1610
Db 1021 CAGCAGCATCAGAGGAGCAGAGTAAACAGGCCCGGACGAGGAGTGAACGAGCCCT 1080
QY 1611 CCTGTCCCTTNGTCCAGGACCAACAGGGCAGAAATCGAGAGTGTCTCTCCCGGC 1670

XX 11-FEB-2000 (first entry)
XX DNA encoding cancer associated antigen KOC-2.
XX Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.
XX Homo sapiens.
XX WO9954738-A1.
XX 28-OCT-1999.
XX 16-MAR-1999; 99WO-US005766.
XX 17-APR-1998; 98US-00061709.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX WPI; 2000-013284/01.
XX Nucleotides representing cancer-associated genes, used to develop
XX products for the diagnosis, monitoring and treatment of cancers.
XX Claim 55; Page 40; 44pp; English.
XX The present sequence represents a cancer associated antigen gene
XX designated KOC-2. The specification also describes a cancer associated
XX antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-
XX 37 melanoma cells. The polypeptide has some homology with MAGS-10,
XX the protein has a repetitive pattern, with repeats rich in serine,
XX proline, glutamine and leucine, and an almost invariable core of the
XX peptide given in AY43877. The CT7 polypeptide can be processed to
XX peptides which provoke lysis by cytolytic T cells. The polynucleotides
XX and polypeptides can be used for treating a cancerous condition and
XX screening for or diagnosing cancerous conditions. The cancer associated
XX antigens can be used as an immunogenic or vaccine composition with an
XX adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony
XX stimulating factor (GM-CSF)
XX SQ Sequence 1708 BP; 447 A; 469 C; 473 G; 314 T; 0 U; 5 Other;
Query Match 85.8%; Score 1669.2; DB 3; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 268 CCAGCAGCCATCATCAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAAGTCT 327
Db 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACCACTGGAGAACCATGCCCTGAAGTCT 89
Qy 328 CCTACATCCCGATGAGCAGATAGCAGGACCTGGAATGGGGCCGAGGGGCTTTG 387
Db 90 CCTACATCCCGATGAGCAGATAGCAGGACCTGGAATGGGGCCGAGGGGCTTTG 149
Qy 388 GCTCTGGGGTACGCCCGCAGGGCTACCTGTGGCAGCGGGGCCCCAGCAAGCAGC 447
Db 150 GCTCTGGGGTACGCCCGCAGGGCTACCTGTGGCAGCGGGGCCCCAGCAAGCAGC 209
Qy 448 AGCAAGTGCATCCCGCTCGGCTCGTGGTCCCAAGTATGTTGGGTGCAATTATG 507
Db 210 AGCAAGTGCATCCCGCTCGGCTCGTGGTCCCAAGTATGTTGGGTGCAATTATG 269
Qy 508 GCAAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGAGTCG 567
Db 270 GCAAGAGGGGGCCACCATCCGCAACATCAAAACAGACCCAGTCCAGATAGAGTCG 329
Qy 568 ATAGGAGAGAACCGAGGTGAGTGAAGAGCCATAGTGTGCACTCCACCCCTGAGG 627
Db 330 ATAGGAGAGAACCGAGGTGAGTGAAGAGCCATAGTGTGCACTCCACCCCTGAGG 389

Qy 628 GCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACCA 687
Db 390 GCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCAATAAGAGGCTAAGGACCA 449
Qy 688 AAACGGGTGACGAGGTTCCCTGGAAGATCTCGGCCATAAATCTTGTAGGGCTCTCA 747
Db 450 AAACGGGTGACGAGGTTCCCTGGAAGATCTCGGCCATAAATCTTGTAGGGCTCTCA 509
Qy 748 TTGGCAAGGAGGACGGAACTTGAAGAGGTAGAGCAAGATACCGAGACAAATAACCA 807
Db 510 TTGGCAAGGAGGACGGAACTTGAAGAGGTAGAGCAAGATACCGAGACAAATAACCA 569
Qy 808 TCTCTCTGTTGCAAGACCTTACCTTTTACAACCTTGAGAGGACCATCTACTGTGAAGGGG 867
Db 570 TCTCTCTGTTGCAAGACCTTACCTTTTACAACCTTGAGAGGACCATCTACTGTGAAGGGG 629
Qy 868 CCATCGAATTTGTTGAGGGCCGAGCAGGAATTAATGAAGAAAGTTCCGGAGGCTATG 927
Db 630 CCATCGAATTTGTTGAGGGCCGAGCAGGAATTAATGAAGAAAGTTCCGGAGGCTATG 689
Qy 928 AGAATGATGTGCTGCCATGAGCTCTACCTGATCCCTGGCCTGAACCTGGCTGTGTAG 987
Db 690 AGAATGATGTGCTGCCATGAGCTCTACCTGATCCCTGGCCTGAACCTGGCTGTGTAG 749
Qy 988 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 1047
Db 750 GTCTTTTCCAGCTTCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 809
Qy 1048 CTCCTATAGTCTCTTTATGAGGCTCCGAGCAGGAGATGTCGAGTGTCTTATCCCGG 1107
Db 810 CTCCTATAGTCTCTTTATGAGGCTCCGAGCAGGAGATGTCGAGTGTCTTATCCCGG 869
Qy 1108 CCCAGCAGTGGGGCCATCATCGCAAGAGGGGAGGAGCAGATCAACAGCTCTCCCGT 1167
Db 870 CCCAGCAGTGGGGCCATCATCGCAAGAGGGGAGGAGCAGATCAACAGCTCTCCCGT 929
Qy 1168 TTCCAGCGCTCCATCAGATTCACACCCGAAACACTGACTCCAAAGTTCTGTATG 1227
Db 930 TTCCAGCGCTCCATCAGATTCACACCCGAAACACTGACTCCAAAGTTCTGTATG 989
Qy 1228 TTATCATCATGACCGCCAGAGCCCAATTCAGGCTCAGGGAGAGATCTATGCGAAAC 1287
Db 990 TTATCATCATGACCGCCAGAGCCCAATTCAGGCTCAGGGAGAGATCTATGCGAAAC 1049
Qy 1288 TCAAGGAGGAGAACTCTTTGTTGCCAAGGAGGAGTGAAGCTGGAGCCACATACGTG 1347
Db 1050 TCAAGGAGGAGAACTCTTTGTTGCCAAGGAGGAGTGAAGCTGGAGCCACATACGTG 1109
Qy 1348 TGCAGCATCAGCAGCTGGCGGGTCAATTTGGCAAGGTGGAAGACCGGTGAACGAGTTGC 1407
Db 1110 TGCAGCATCAGCAGCTGGCGGGTCAATTTGGCAAGGTGGAAGACCGGTGAACGAGTTGC 1169
Qy 1408 AGAATTTGACGGCAGCTGAGTGTAGTACCAAGAGCAGACCCCTCATGAGAACGACC 1467
Db 1170 AGAATTTGACGGCAGCTGAGTGTAGTACCAAGAGCAGACCCCTCATGAGAACGACC 1229
Qy 1468 AGGTTCATCGTGAATTCATCGACATTTCTATGCGAGTCAAGTGGCTCAACGGAAGTCC 1527
Db 1230 AGGTTCATCGTGAATTCATCGACATTTCTATGCGAGTCAAGTGGCTCAACGGAAGTCC 1289
Qy 1528 GAGACATCCTCGCCGAGGTTAAGCAGCAGCATCAGAGGAGCAGATTAACCGGCCAGG 1587
Db 1290 GAGACATCCTGGCCCGAGGTTAAGCAGCAGCATCAGAGGAGCAGATTAACCGGCCAGG 1349
Qy 1588 CACGAGGAGATGACACCGCCCTCCCTGTCCTTNGAGTCCAGGACCAACCGGCGAGAA 1647
Db 1350 CACGAGGAGATGACACCGCCCTCCCTGTCCTTNGAGTCCAGGACCAACCGGCGAGAA 1409
Qy 1648 ATCGAGAGTGTCTCTCCCGGAGGCTGAGATGAGTGGGAATCCGGGACACNTGGGC 1707
Db 1410 ATCGAGAGTGTCTCTCCCGGAGGCTGAGATGAGTGGGAATCCGGGACACNTGGGC 1469
Qy 1708 CGGGCTGTAGATCAGGTTTGGCCCATCTTGATGAGAAAGATGTTCCAGTGGAGAACCTTGA 1767

CC	macrophage-colony stimulating factor (GM-CSF)	
XX		
SQ	Sequence 1946 BP; 502 A; 528 C; 553 G; 358 T; 0 U; 5 Other;	
	Query Match 99.7%; Score 1941; DB 3; Length 1946;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCTGTAGCGAGGGCTGGGGGCTGCTCTGTCTGCCCTTCTTGGCGGCTGCGGCTCTCAGC	60
DB	1 GCTGTAGCGAGGGGCTGGGGGCTGCTCTGTCTGCCCTTCTTGGCGGCTGCGGCTCTCAGC	60
QY	61 CCACCCAGAGGCGGGGTGGAGGGCGAGTGTCTAGCTTCCCGGGTTAGAGGCCGGAANA	120
DB	61 CCACCCAGAGGCGGGGTGGAGGGCGAGTGTCTAGCTTCCCGGGTTAGAGGCCGGAANA	120
QY	121 TTCAAAATCCGAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTTCTGGCTC	180
DB	121 TTCAAAATCCGAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCTTCTGGCTC	180
QY	181 AGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCCGAGAGTGAAGCGGCAGTGTGA	240
DB	181 AGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACCCGAGAGTGAAGCGGCAGTGTGA	240
QY	241 ATGTACCTTATTCCAAACCGGGAGCAGACCGCAAGCCATCTGAAGCTGAATGGCCACC	300
DB	241 ATGTACCTTATTCCAAACCGGGAGCAGACCGCAAGCCATCTGAAGCTGAATGGCCACC	300
QY	301 AGTTGAGAACCATGCCCTCAAGTCTCTTACATCCCGGATGAGCAGATGACACAGGGAC	360
DB	301 AGTTGAGAACCATGCCCTCAAGTCTCTTACATCCCGGATGAGCAGATGACACAGGGAC	360
QY	361 CTGAGAATGGCGCGCAGGGGGCTTTGGCTCTCGGGTTCAGCCCGCCAGGGGTCACTTG	420
DB	361 CTGAGAATGGCGCGCAGGGGGCTTTGGCTCTCGGGTTCAGCCCGCCAGGGGTCACTTG	420
QY	421 TGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCTCGGTGC	480
DB	421 TGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCCCTTCGGCTCTCGGTGC	480
QY	481 CCACCCAGTAGTGGGTGCCATTATTGSCAAGGAGGGGCCACCATCCGCAACATCAAA	540
DB	481 CCACCCAGTAGTGGGTGCCATTATTGSCAAGGAGGGGCCACCATCCGCAACATCAAA	540
QY	541 AACAGACCCAGTCCCAAGTAGACGTGCATAGGAAGAGAGAACCGAGGTGCAGCTGAAAAG	600
DB	541 AACAGACCCAGTCCCAAGTAGACGTGCATAGGAAGAGAGAACCGAGGTGCAGCTGAAAAG	600
QY	601 CCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTTGAAGATGATCTTGGAGA	660
DB	601 CCATCAGTGTGCACTCCACCCCTGAGGGCTGCTCTCCGCTTGAAGATGATCTTGGAGA	660
QY	661 TTATGCAATAAGAGGCTAAGACACCAAAACGGCTGACAGGTTCCCTGAGATCCTCTGG	720
DB	661 TTATGCAATAAGAGGCTAAGACACCAAAACGGCTGACAGGTTCCCTGAGATCCTCTGG	720
QY	721 CCCATAAATCTTTTAGGGCTCTCATTTGGCAAGGAAGAGCGGAACCTGAAGAGGTAG	780
DB	721 CCCATAAATCTTTTAGGGCTCTCATTTGGCAAGGAAGAGCGGAACCTGAAGAGGTAG	780
QY	781 AGCAAGATACCGAGACAAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACAACC	840
DB	781 AGCAAGATACCGAGACAAAAATCACCATCTCTCGTTGCAAGACCTTACCTTTACAACC	840
QY	841 CTGAGGAGCACCATCACTGTGAAGGGGCCATCGAATTTGTCAGGGCCGAGCAGGAAA	900
DB	841 CTGAGGAGCACCATCACTGTGAAGGGGCCATCGAATTTGTCAGGGCCGAGCAGGAAA	900
QY	901 TAATGAAGAAAGTTGGGAGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGA	960
DB	901 TAATGAAGAAAGTTGGGAGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGA	960
QY	961 TCCCTGGCCGTGAACCTGGCTGCTGTAGTCTTTCCAGCTTCATCCAGCGCAGTCCCGC	1020
DB	961 TCCCTGGCCGTGAACCTGGCTGCTGTAGTCTTTCCAGCTTCATCCAGCGCAGTCCCGC	1020
QY	1021 CGCTCCAGCAGCGTTACTGGGGTGTCTCTATAGTCTCTTTATGAGGCTCCCGAGC	1080
DB	1021 CGCTCCAGCAGCGTTACTGGGGTGTCTCTATAGTCTCTTTATGAGGCTCCCGAGC	1080
QY	1081 AGGATGTTGTCAGGTGTTTATCCCGCCAGGCGAGTGGGCCCATCATCGGCAAGAAG	1140
DB	1081 AGGATGTTGTCAGGTGTTTATCCCGCCAGGCGAGTGGGCCCATCATCGGCAAGAAG	1140
QY	1141 GGCAGCACATCAAAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATGACACCCCG	1200
DB	1141 GGCAGCACATCAAAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATGACACCCCG	1200
QY	1201 AAACACCTGACTCCAAAGTTCTATGCTGATGTTATCATCTGGACCGCAGAGGCCAATCA	1260
DB	1201 AAACACCTGACTCCAAAGTTCTATGCTGATGTTATCATCTGGACCGCAGAGGCCAATCA	1260
QY	1261 AGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAACTTTTGGTCCCAAGAGG	1320
DB	1261 AGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAACTTTTGGTCCCAAGAGG	1320
QY	1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGTCATTTGCA	1380
DB	1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCGGTCATTTGCA	1380
QY	1381 AAGTGGAAAAACGGTGAACGAGTTGACAGATTTGACGGCAGCTGAGTGGTAGTACCAA	1440
DB	1381 AAGTGGAAAAACGGTGAACGAGTTGACAGATTTGACGGCAGCTGAGTGGTAGTACCAA	1440
QY	1441 GAGACACAGACCCCTGATGAGAACGACACAGCTCATCTGAAAAATCATCGGACATTTCTATG	1500
DB	1441 GAGACACAGACCCCTGATGAGAACGACACAGCTCATCTGAAAAATCATCGGACATTTCTATG	1500
QY	1501 CCAAGTGAAGTGGTCAACCGGAAGATCCGAGACATCTGGGCCAGGTTAAGCAGCAGATC	1560
DB	1501 CCAAGTGAAGTGGTCAACCGGAAGATCCGAGACATCTGGGCCAGGTTAAGCAGCAGATC	1560
QY	1561 AGAAGGACAGAGTAAACAGGCCCGCAGGACGAGGAGTCAACAGCCCTCCCTGTCCCT	1620
DB	1561 AGAAGGACAGAGTAAACAGGCCCGCAGGACGAGGAGTCAACAGCCCTCCCTGTCCCT	1620
QY	1621 TNGAGTCCAGACAAACAAACGGGCGAGAAATCGAGTGTGCTCTCCCGGAGGCTCAGAA	1680
DB	1621 TNGAGTCCAGACAAACAAACGGGCGAGAAATCGAGTGTGCTCTCCCGGAGGCTCAGAA	1680
QY	1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA	1740
DB	1681 ATGAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGGCCACTTGATGA	1740
QY	1741 GAAAGATTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCAACCCCAATTTGGGCC	1800
DB	1741 GAAAGATTTCCAGTGAAGAACCTGATCTNTCAGCCCAACCAACCCCAATTTGGGCC	1800
QY	1801 AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGACCTTTTAAACGTGGAT	1860
DB	1801 AACACTGNTGCCCTCGGGGTGTGAGAAATTTAGCGCAAGGACCTTTTAAACGTGGAT	1860
QY	1861 TGTTTAAAGAAAGCTCTCCAGGCCCAACCAAGAGGGTGGATCACACCTCAGTGGGAGAAA	1920
DB	1861 TGTTTAAAGAAAGCTCTCCAGGCCCAACCAAGAGGGTGGATCACACCTCAGTGGGAGAAA	1920
QY	1921 AATTAATTTCTTCAGGTTTTTAAAA 1946	
DB	1921 AATTAATTTCTTCAGGTTTTTAAAA 1946	
RESULT 2		
AAZ36151		
ID AAZ36151 standard; DNA; 1708 BP.		
XX		
AC AAZ36151;		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 03:34:49 ; Search time 719.632 Seconds
(without alignments)
11503.795 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtagcggaggctggg.....atttccttcagggttttaaaa 1946

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2003as: *
8: geneseqn2003bs: *
9: geneseqn2003cs: *
10: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1941	99.7	1946	3	Aaz36153 An altern
2	1699.2	85.8	1708	3	Aaz36151 DNA encod
3	1389	71.4	2780	4	Aak94782 Human ful
4	1297.8	66.7	2224	2	Aaz10617 cDNA enco
5	1085.4	55.8	1182	4	Aas26148 Human cDN
6	1085.4	55.8	1182	7	Abx73489 Human nov
7	1069.8	55.0	1129	7	Aas26566 Human cDN
8	1069.8	55.0	1129	7	Abx73907 Human nov
9	732.8	37.7	833	4	Aak91969 Human cDN
10	732.8	37.7	833	4	Aak93655 Human cDN
11	674.2	34.6	1740	3	Aac66035 Human lun
12	674.2	34.6	1740	6	AbL49254 Human lun
13	674.2	34.6	1740	6	AbQ92440 Human lun
14	674.2	34.6	1743	6	AbL49299 Human lun
15	674.2	34.6	1743	6	AbL49297 Human lun
16	674.2	34.6	1743	6	AbQ92483 Human lun
17	674.2	34.6	1743	6	AbQ92485 Human lun
18	674.2	34.6	1743	8	ADA28540 Recombina
19	674.2	34.6	1743	8	ADA28537 Recombina
20	674.2	34.6	1764	6	AbL49283 Human lun
21	674.2	34.6	1764	6	AbQ92469 Human lun
22	674.2	34.6	1764	8	ADA28518 Recombina
23	672.6	34.6	1740	8	ADA28437 Human lun

24	672.6	34.6	4159	3	Aaz36150 DNA encod
25	672.6	34.6	4181	3	Aac65900 Human lun
26	672.6	34.6	4181	6	AbL65405 Lung canc
27	672.6	34.6	4181	6	AbL49119 Human lun
28	672.6	34.6	4181	6	AbQ92305 Human lun
29	672.6	34.6	4181	8	ADA28285 Human lun
30	672.6	34.6	4181	9	ADD14661 Human src
31	672.6	34.6	4181	9	ADe53470 Human lun
32	672.6	34.6	4434	8	ACH03940 Human cDN
33	672.6	34.6	4601	6	ABA99958 Human KOC
34	620.8	31.9	1985	5	Aas76779 DNA encod
35	602.6	31.0	4264	5	Aas866150 DNA encod
36	585.4	30.1	3283	3	Aaz36154 An altern
37	585.4	30.1	3667	5	Aas70982 DNA encod
38	585.4	30.1	3667	6	AbS76442 cDNA enco
39	570.6	29.3	2010	5	Aas70981 DNA encod
40	570.6	29.3	3412	3	Aaz36152 DNA encod
41	570.6	29.3	3694	4	Aas26150 Human cDN
42	570.6	29.3	3694	7	Abx73491 Human nov
43	532.2	27.3	1707	7	ACA90176 cDNA enco
44	267.2	13.7	2290	4	Aas26153 Human cDN
45	267.2	13.7	2290	7	Abx73494 Human nov

ALIGNMENTS

RESULT 1
Aaz36153
ID Aaz36153 standard; DNA; 1946 BP.
XX
AC Aaz36153;
XX
DT 11-FEB-2000 (first entry)
XX
DE An alternative form of DNA encoding cancer associated antigen KOC-2.
XX
KW Cancer associated antigen; KOC-2; cancer; vaccine; CT7; ss.
XX
OS Homo sapiens.
XX
PN WO9954738-Al.
XX
PD 28-OCT-1999.
XX
PF 16-MAR-1999; 99WO-US005766.
XX
PR 17-APR-1998; 98US-00061709.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX
DR WPI; 2000-013284/01.
XX
PT Nucleotides representing cancer-associated genes, used to develop
XX products for the diagnosis, monitoring and treatment of cancers.
XX
PS Claim 55; Page 42; 44pp; English.
XX
CC The present sequence represents an alternative form of a cancer
XX associated antigen gene designated KOC-2. The specification also
XX describes a cancer associated antigen designated CT7. The CT7
XX polynucleotide was isolated from SK-MEL-37 melanoma cells. The
XX polypeptide has some homology with MAGE-10, limited to about 210 carboxy
XX terminal amino acids. The amino terminal of the protein has a repetitive
XX pattern, with repeats rich in serine, proline, glutamine and leucine, and
XX an almost invariable core of the peptide given in AAY43877. The CT7
XX polypeptide can be processed to peptides which provoke lysis by cytolytic
XX T cells. The polynucleotides and polypeptides can be used for treating a
XX cancerous condition and screening for or diagnosing cancerous conditions.
XX The cancer associated antigens can be used as an immunogenic or vaccine
XX composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte

This Page Blank (uspto)

Db 903 |||||ATGATCTTGAGATCATGCAAGAGGAGGCAAGACACGAGCAGCTGATGAGTGCCT 962
Qy 708 CTGAGATCTGGCCCAATAAATCTTTGTAGGGCTCTCATTTGGCAAGGAGGACGGAAC 767
Db 963 CTGAAATCTTGGCCCAATAAATCTTTGTAGGGCTCTCATTTGGCAAGGAGGCGGAAC 1022
Qy 768 CTGAAGAAGGTAGAGCAAGATACCGAGACAAAATACCATCTCTCTCGTTGCAAGACCTT 827
Db 1023 TTGAGAAAGTGGAGCAGGATACTGAGACAAAATACCATCTCATCTTGCAGGACCTG 1082
Qy 828 ACCCTTTACACCTGAGAGACCATCATCTGTGAAGGGGCCATCGAGAAATTTGTCAGG 887
Db 1083 ACCCTGTACAAACCCGAAAGAACGATCACAGTGAAGGGCTCTATCGAGAACTGCTGCAA 1142
Qy 888 GCCGAGCAGGAATAAATGAGAAAGTTCCGGAGGCCCTATGAGAAATGATGTGGCTGCCATG 947
Db 1143 GCAGAGCAGGAGATCATGAGAAAGTGAAGAGCCCTACGAGAACGATTTGACGCCATG 1202
Qy 948 AGC-----TCTACCTGATCCCTGGCTGAACTGGCTGTGTAGGTCTTTTCCAGCT 1001
Db 1203 AGCCTGCAATCTCATCTCATCTCCCTGGCTCAACCTGGCTGCAGTTGGCTCTTCCCTGCC 1262
Qy 1002 TCATCCAGCGAGTCCCGCCCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTCC 1061
Db 1263 TCCTCCAATGCACTACCTCTCCAGCAGCGTCTCCGGGGTGTCTCCATACAGCTCC 1322
Qy 1062 TTTATGAGGCTCCCGAGCAGGAGATGTTGAGGTGTTTATCCCGCCAGGCGAGTGGC 1121
Db 1323 TTCAATG---CTCCGAGCAGGAGACCGTACACGTCTTCATCTCTGCCAGCGGTTGTT 1379
Qy 1122 GCCATCATCGGCAAGAGGGCAGCACAATCAACAGCTCTCCCGGTTTGGCAGCGCTCC 1181
Db 1380 GCCATCATTTGCAAGAGGGCCAGCACAATCAAGCAGCTCTCCCGGTTTGGCAAGTGCCTCT 1439
Qy 1182 ATCAAGATTGACACCGGAAACCTGACTCCAAAGTTGATGTTATCATCACTGGA 1241
Db 1440 ATTAAGATTGACCCCGGAGACCGGACTCCAAAGTGCAGTGGTGGTCAACCGGC 1499
Qy 1242 CCGCCAGAGGCCCAATTCAGGCTCAGGGAAGATCTATGCAAACTCAAGGAGGAGAAC 1301
Db 1500 CCTCCAGAGCTCAGTTCAAGGCAACAGGCGAGTTTATGGAGCTGAAGGAGGAGAAC 1559
Qy 1302 TTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
Db 1560 TTCTTTGGGCCCAAGGAGGAAGTGAAGCTGGAGACGCACATCCGTTGCCCGCTCGGCT 1619
Qy 1362 GCTGCGGGCTCATTTGGCAAGGTGGAAAAACGGTGAACGAGTTGCAGAAATTTGACGCA 1421
Db 1620 GCAGGAGGGTGAATCGGCAAGGGGGCAAAACCCGTCAATGAGCTCAGAACCTGACGGCT 1679
Qy 1422 GCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACGAGTCACTCGTAAA 1481
Db 1680 GCAGAGGTGGTGTTCACGGGATCAGACCCCTGATGAGATGAGCAGTCAATGTGAAG 1739
Qy 1482 ATCATCGGACATTTCTATGCCAGTCAGATGGCTCAACCGAAGATCCGAGACATCTGGCC 1541
Db 1740 ATCATCGGCACTTCTATGCCAGCCAGATGGCAGCGCAAAATCCGGGACATCTGGCC 1799
Qy 1542 CAGGTTAAGCAGCAGCATCAGNAGGACAGATACCGGCCCGGAGGAGGAGTGA 1601
Db 1800 CAGGTGAAGCAGCAGCACCAGAAAGGACAGCGGCCAGCTGCAAGCGAGGAGGAATGA 1859

Db 1777 GCCGAGGTGTCGTGCCGAGAGACCAGACCCAGACAGAGAACAGCAAGTCAATGTTAAG 1836
QY 1482 ATCATCGGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTTGCC 1541
Db 1837 ATCATCGGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCCTTGCC 1896
QY 1542 CAGGTTAAGCAGCAGCATCAGAACGGGACAGAGTAACAGCCAGCCAGGACCGAGGAGTGA 1601
Db 1897 CAAGTTAAGCAGCAGCATCAGAACGGGACAGAGTAACAGCCAGGACCGAGGAGTGA 1956
QY 1602 -CCAGGCCCTCCCTGTCCTCTTNGAGTCCAGGACAAACAGCGGACGAG----- 1646
Db 1957 CCCCAGCCTTCTCTGTCCTCCATGGCTCCAGATCAGCAGGCGGACACACAGGACTGGAGGAC 2016
QY 1647 -----ATCAGAGTGTGCTCTCCCGGAGGCTGAGATGAGTGGGAATCCGGGACAC 1701
Db 2017 GGGTGAGGGCCGCGGTCTCTCTAGCAGGCTGAGATGAGTGGGAATCCGGGACAC 2075
QY 1702 NTGGGCGGGCTGTAGATCAGGTTTGCCCA 1731
Db 2076 TTGGGTCTGGCTGGAGATCAGGTTTGACAA 2105

RESULT 15
AF026527
LOCUS 2021 bp mRNA linear VRT 30-OCT-1997
DEFINITION Gallus gallus zipcode-binding protein (ZBP1) mRNA, complete cds.
ACCESSION AF026527
VERSION AF026527.1 GI:2570920
KEYWORDS
SOURCE
ORGANISM Gallus gallus (chicken)
Bukaryofa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2021)
AUTHORS Ross,A.F., Oleynikov,Y., Kislauskis,E.H., Taneja,K.L. and Singer,R.H.
Characterization of a beta-actin mRNA zipcode-binding protein
Mol. Cell. Biol. 17 (4), 2158-2165 (1997)
97220007
9121465
REFERENCE 2 (bases 1 to 2021)
AUTHORS Oleynikov,Y.S., Ross,A.F. and Singer,R.H.
Direct Submission
TITLE Submitted (24-SEP-1997) Cell Biology, Albert Einstein College of
JOURNAL Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA
FEATURES
Location/Qualifiers
1..2021
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/cell_type="fibroblast"
/dev_stage="embryo"
1..2021
/gene="ZBP1"
129..1859
/gene="ZBP1"
/function="binds a zipcode element in beta-actin mRNA
3'UTR and is involved in its localization"
/notes="beta-actin mRNA zipcode-binding protein"
/codon_start=1
/product="zipcode-binding protein"
/protein_id="AAB82295.1"
/db_xref="GI:2570921"
/translation="MNKLVIGNESVTPADLEKVFNDHKISFGSGLVKSGYAFVDC
PQEWAKAIEFTSGVLEHQLIEHSVPKQSRKTIQIRNPQLRWEVLGGLA
QYGVENCQVNTDSYAVNVYTNREOTROAIMKLNHLENHVLKVSYPDEQSV
QCPENRGSGFARGAPRGSPVTAGAPVKQOPVDIPLRLVPTOVVGAIGKEGATI
RNITKQTSKIDVHRKENAGAEKALISHTPEGCSAAKMLEIMQKEAKDYKTADE
VFLKILAHNFVGLRIGKGRNLKVEQDTETKITISSIDLDLYNPRTITVKGSI
NCKAQQEIMKKVREYFNDVAMSLQSHLIPGLNLAAYGLFPASSNAPVPPPPSVSG
AAPYSFMPPEQETHVFIPAQAVGAILGKGGHILKQLSRFASASIKIAPPETPDSKV

ORIGIN

Query Match 53.8%; Score 1046; DB 5; Length 2021;
Best Local Similarity 82.1%; Pred. No. 8.3e-246;
Matches 1231; Conservative 0; Mismatches 260; Indels 9; Gaps 2;
QY 108 AGGAGCCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTACTGGAC 167
Db 363 AGGAGCCGGAAATTCAAATCCGAAATATCCACCCAGCTCCGATGGGAAGTACTGGAT 422
QY 168 AGCTCTGCTGGCTCAGTATGTTAGTAGAAGTGTGAGCAAGTGAACACCGAGAGTGAG 227
Db 423 GGCTTCTAGCAATATGTCATGTAGAAATGTGAGCAAGTGAACACAGACAGTGAAG 482
QY 228 ACGGCAGTGTGAAATGTCACTTATCCACCGGAGACAGACCAAGCAAGCATCATGAAG 287
Db 483 ACCGCTGTGGTAAATGTCACTTATCCACCGGAGACAGACCAAGCAAGCATCATGAAG 542
QY 288 CTGAATGGCCACAGTTGGAGAACCATGCCCTGAAGGTCTCTACATCCCGATGAGCAG 347
Db 543 TTAAACGGGACACAGCTGGAGAACCACTGCTGTGAAGTCTCTACATCCCTGATGAGCAG 602
QY 348 ATAGCAGAGGACCTGAGAAATGGGCGCCGAGGGGGCTTTGGCTCTCGGGGTTCAGCCCGC 407
Db 603 TCCGTGACGGGCCAGAGAAATGGGCGGGGTGGCTTTGGGCGCCGAGGTGCCCGCCCG 662
QY 408 CAGGGCTCAGCTGTGGCAGCGGGGGCCCGAGCCAGCAGCAGCAAGTGAACATCCCGCTT 467
Db 663 CAGGGCTCCCTGTCTCACTGAGGGGTCCAGTCAAGCAGCAGCCCGTGGACATCCCGCTC 722
QY 468 CGGCTCTCGTGGTCCCAACCCAGTATGTGGGTGCCATTTATTGGCAAGAGGGGGCCACCATC 527
Db 723 CGTCTGCTGGTGGCCACGAGTATGTGGGTGCCATCATTTGGCAAGAGGGGGCCACCATC 782
QY 528 CGCAACATCAGAAAAAGACCCAGTCCAGATAGACGTGTCATAGGAAGAGAGAACGCAGT 587
Db 783 AGGAACATTACCAAGCAGACAGTCCAAGATTGATGTGCAACCGTAAGAGAAATGCAGGA 842
QY 588 GCAGCTGAAAAAGCCATCAGTGTGCACCTCCAGGGGTGCTCTCCCGCTGTGAAG 647
Db 843 GCTGCAGAAAAAGTATCAGATCCACTCCACCCCGAGGGGTGCTCTCTGCTGCCTGCAAA 902
QY 648 ATGATCTTGGAGATTATGCATAAAGAGGCTAAGGACACCAAAACCGCTGACGAGGTGCC 707

RESULT 14
AF541940 2105 bp mRNA linear ROD 25-NOV-2003
LOCUS Rattus norvegicus b-actin zipcode binding protein 1 mRNA, complete cds.
DEFINITION
AF541940
VERSION AF541940.1 GI:27464837
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE
1 (bases 1 to 2105)
Eom, T., Antar, L.N., Singer, R.H. and Bassell, G.J.
Localization of a beta-actin messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses
J. Neurosci. 23 (32), 10433-10444 (2003)
MEDLINE 22979172
PUBMED 14614102
REFERENCE
2 (bases 1 to 2105)
Eom, T., Singer, R.H. and Bassell, G.J.
Direct Submission
TITLE Submitted (27-AUG-2002) Neuroscience, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..2105
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
223..1956
/note="rZBP1"
/codon_start=1
/product="b-actin zipcode binding protein 1"
/protein_id="AA016210.1"
/db_xref="GI:27464838"
/translation="MNKLYIGNLNESTVPADLEKVPFAEHKLSYGOFLVKSGYAFVDC
PDEHWMKAIETFSQKVELOKLEIHSVPKORSRKIOIRNIPPOLRWVLSLLA
QYGTVENCEQNTSETAVNVTYSNREQTQAIKMLNGHOLENHAKVSVIPDEQIA
QGPENRRGGSGDPQPGFVAGAPAKQQQVDIPURLVFPYVGYALIGKEGATI
RNIQTQSKDIDVHRKENAGAAEKAISVHSPEGSACKMILEMHKEADTKTAD
VPLKILAHNNFVGRILIGENRLAKVEQDTETKITISSLODLTYNPERTIVKGAIE
NCCRAEQIIMKKVREAYENDVAAMSLOSHLIPGLNLAAGLFPASSAVPPPPSVTG
AAPYGSFWQAPQEMVOVEIQAQAVGALIGKKGOHIKOLSRFASAIKIAPETPDSK
VRMVITGPQAPQAKQRIYIGKLEENFPFKBEVKLETHIRVPASAGSVIGKGGK
TVNELQNTAAEVVVPVDPQTPDENDQVIVKLIIGHFYASQMAQRKIRDLIAQVKQHQK
GQSNQAQARRK"
ORIGIN
Query Match 64.2%; Score 1250; DB 10; Length 2105;
Best Local Similarity 86.7%; Pred. No. 6.4e-296;
Matches 1431; Conservative 0; Mismatches 192; Indels 27; Gaps 4;
QY 108 AGGAGCGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167
DB 457 AGGAGTCGGAAATACAGTCCGGATATTCACCTCAGCTCCGATGGGAAGTCTGGAT 516
QY 168 AGCCTGCTGGCTCAGTATGGTACAGTAGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 227
DB 517 AGCCTTCTGGCCAGTACGGCAGTGGAGAACTGTGAGCAAGTGAACACCGAGAGTGAG 576
QY 228 ACGGAGTGGTGAATGTACCTTATTCACCGGGAGCAGACCGAGCAAGCCATCATGAAG 287
DB 577 ACGGCGGTGGTCAACGTACCTTATCTAACCGGGAGCAGACCGAGCAAGCCATCATGAAG 636
QY 288 CTGAATGGCCACAGTTGGAGAACCATGCCCTGAAAGGTTCTCTACATCCCCGATGAGCAG 347
DB 637 CTAATATGGCCATCAACTGGAGAACCATGCCCTGAAAGGTTCTCTACATACCTCATGAGCAG 696
QY 348 ATAGCACAGGACCTGAGAAATGGGCGCGAGGGGGCTTTGGCTCTCGGGGTTCAGCCCCG 407

Db 697 ATAGCACAGGTCCTGAGAAATGGGCGCGGTGGAGGCTTTGGTCTCGAGGGCAGCCAGG 756
QY 408 CAGGGCTCACCTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCCCTT 467
Db 757 CAGGGTCAACGGTAGCAGCAGGAGGGCGCCAGCCAGCAGCAGCAGTGGACATCCCTCTC 816
QY 468 CGGCTCCTGTGGTCCACCCAGTATGTGGGTGTCCTATTATTGGCAGGAGGGGCCACCATC 527
Db 817 CGGCTCCTGTGGTCCACCCAGTATGTGGGTGTCCTATTATTGGCAGGAGGGGCCACCATC 876
QY 528 CGCAACATCACAAACAGACCCAGTCCAAAGATAGACGTGCATAGAGAGGAGAACGACAGT 587
Db 877 CGAAACATCACAAACAGACCTCAGTCCAAAGATAGACGTGCACAGGAGGAGAACGCGGG 936
QY 588 GCAGCTGAAAAGGCATCAGTGTGCATCCACCCTGAGGCTGTCTCTCGCTTTGTAAAG 647
Db 937 GCTCGGAGAGAGCCATCAGTGTGCATCCAGCCCTGAAAGCTGTCTCTCGGCTGTAAAG 996
QY 648 ATGATCTTGGAGATTATGCATAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTCCT 707
Db 997 ATGATCTTGGAGTTATGCACAAGGAGGCAAGGACACCAAGACCGCAGACGAGATTCTC 1056
QY 708 CTGAAGATCTCGGCCCATATAAATCTTGTAGGGGCTCTCATTTGGCAAGGAGAACGGAAC 767
Db 1057 CTGAAGATCTCGGCTCATATAAATCTTGTGGGGGACCTCATTTGGAAGGAGGCGGAAC 1116
QY 768 CTGAAGAGGTAGAGCAAGATACCGAGACAAAATACCATCTCTCTGTTGGAAGACCTT 827
Db 1117 TTGAAGAAAGTGGAGCAGGATACAGAGACTAAGATCACCATCTCTCTCGCTCCAGGATCTC 1176
QY 828 ACCCTTTCACCCCTGAGAGACCATCATCTGTAGGGGGCCATCGAGAAATTTGTCAGG 887
Db 1177 ACGCTCTATAACCCCTGAAAGACCATATACTGTGAAGGGGGCCATTTGAGAACTGCTGCG 1236
QY 888 GCCGAGCAGGAAATAATGAAGAAAGTTCGGGAGGCTATGAGAAATGATGTGCTGCCATG 947
Db 1237 GCTGAGCAGGAGATCATGAGAAAGTTTCGAGAGGCTTACGAGAAATGACGTGCTGCCATG 1296
QY 948 AGC-----TCTCACTCTGATCCCTGGCTGAACTCCCTGGCTGTGTAGTCTTTTCCAGCT 1001
Db 1297 AGCTTGCAGTCTCACTCTCATCCCTGCTCAACCTGGCTGTGTGCTGCTCTCTTCCAGCT 1356
QY 1002 TCATCCAGCAGTCCCGCCGCTCCAGCAGCCTTACTGGGCTGCTCCCTATAGCTCC 1061
Db 1357 TCTTCTAGTGGCTTCCCGCTCTCCAGCAGCCTCCTCCAGCAGCTGCTGCTGCTGCTCC 1416
QY 1062 TTTATGCAAGTCCCGAGCAGGAGATGCTGCAAGTGTGTTATCCCGCCCGCAGGAGTGGC 1121
Db 1417 TTTATGCAAGTCCCGAGCAGGAGATGTTACAGTGTGTTATCCCGCCCGCAGGCTGTGGC 1476
QY 1122 GCCATCATCGGCAAGAGGGCAGCAGATCAAAAGCTCTCCCGTTCCTGAGCGGCTCC 1181
Db 1477 GCCATCATCGGCAAGAGGGCAGCAGATCAAAAGCTCTCCCGTTCCTGAGCGGCTCC 1536
QY 1182 ATCAAGATTGACACCGAAACACCTGACTCCAAAGTTCGTATGTTATCATCTACTGA 1241
Db 1537 ATCAAGATTGCTCCACCAAGAACACCTGACTCCAAAGTTCGAATGTTGTCTACTGA 1596
QY 1242 CGCCAGAGGGCCCAATTCAGAGGCTCAGGAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1301
Db 1597 CCCCAGAGGCTCAGTTCAAGGCTCAGGAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1656
QY 1302 TTTCTTTGTTCCAAAGAGAGTGAAGCTGAGAGACCCACATPACGTGTGCCAGCATCAGCA 1361
Db 1657 TTTCTTTGTTCCAAAGAGAGTGAAGCTGAGAGACCCACATCCGGGTTCAGCTTTCAGCA 1716
QY 1362 GCTGCGCGGCTCATTTGGCAAGGTCGAAACCGTGAACGAGTTCAGAAATTTGACGCA 1421
Db 1717 GCTGCGCGGCTCATTTGGCAAGGTCGAAACCGTGAACGAGTTCAGAAATTTGACGCA 1776
QY 1422 GCTGAGGTGGTAGTACCAAGAGACAGACCCCTGATGAGAACGACGAGTCTATCGTGA 1481

QY	1302	TTCTTTGGTCCCAAGAGGAAAGTGAAGCTGGAGAGCCACATACGTGTGGCAGCATCAGCA	1361
Db	1565	TTCTTTGGTCCCAAGAGGAAAGTGAAGCTGGAGAGCCACATACGTGTGGCAGCATCAGCA	1624
QY	1362	GCTGGCCGGGTCTATGCGAAGGTGGAAACCGTGAACGAGTTGCAGAAATTTGACGCGCA	1421
Db	1625	GCGGGCCGGGTCTATGCGAAGGTGGAAACCGTGAACGAGTTGCAGAAATTTGACGCGCA	1684
QY	1422	GCTGAGGTGGTGTAGTACCAAGAGACGACACCCCTGATGAGAACGACACCGAGTTCATCGTGA	1481
Db	1685	GCTGAGGTGGTGTAGTACCAAGAGACGACACCCCTGATGAGAACGACACCGAGTTCATCGTGA	1744
QY	1482	ATCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACGAGAGATCCAGAGATCTCGTGGC	1541
Db	1745	ATCATCGGACATTTCTATGCGAGTCAAGTGGCTCAACGAGAGATCCAGAGATCTCGTGGC	1804
QY	1542	CAGTTTAAGCAGCAGCATCAGAGGGACAGAGTAAACGAGCCCGAGGACGAGGAGAGTGA	1601
Db	1805	CAAGTTTAAGCAACAGCAGCAGAGGGACAGAGTAAACGAGCCCGAGGACGAGGAGAGTGA	1864
QY	1602	-CCAGCCCTCCCTCTGCTCCCTTNGAGTCCAGGACAAACACCGGCGAGAA-----	1647
Db	1865	CCCGCCCTCCCTCTGCTCCCTTNGAGTCCAGGACAAACACCGGCGAGAA-----	1924
QY	1648	-----ATCGAGAGTGTGCTCTCCCGGCGAGCCTGAGAAATGAGTGGGAATCCGGGAC	1699
Db	1925	GCGGGTGGAGGGCGCGTGTGTTTTTCCAGCAGGCTGAGAAATGAGTGGGAATCAGGG-C	1983
QY	1700	ACNTGGGCGGGCTGTAGATCAGGTTTCCCTGATGTTGAGAAAGATGTTCCAGTGGG	1759
Db	1984	ATTGCGGCTGGCTGGAGATCAGGTTTGCACACTGTATTGAGAAACAATGTTCCAGTGGG	2043
QY	1760	AACCTCATCTNTCAGCCCAACACCCCACTTGGCCCAACACCTGTTTCCCTCGG	1819
Db	2044	AATCTCATCTCTGCCCCCAA--TTGAGCCAGCTGGCCACAGCCCTTGGATAT	2101
QY	1820	GGTGTCAAGAAATNTAGCGCAAGGACATTTTAAACGTGGATGTTTAAAGAGCTCTCCA	1879
Db	2102	CACCATTTGCAATCATAGCTTGGTGTGCTTTTAAACGTGGATGTTT--TGAAGTTCTCCA	2159
QY	1880	GGCCCCCAGAGGGTGGATCACACTCTCAGTGGGAGAGAAATAAATTTCTCTCAGT	1939
Db	2160	GCCTCCATGGA-AGATGGGTGAGATCCCATGGGAGAGAAATAAATTTCTCTCAGT	2218
QY	1940	TTTA 1943	
Db	2219	TTTA 2222	
RESULT 13			
LOCUS	BC051679	2444 bp mRNA linear ROD 08-OCT-2003	
DEFINITION	Mus musculus insulin-like growth factor 2, binding protein 1, mRNA	(cDNA clone MGC:60613 IMAGE:30008106), complete cds.	
ACCESSION	BC051679.1	GI:30354043	
VERSION	MGC.		
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 2444)		
AUTHORS	Stausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,		

Vallalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Helton, E., Rettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A.C., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257
12477932
2 (bases 1 to 2444)
Strausberg, R.
Direct Submission
Submitted (30-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Kenneth R. Boheler (National Institute on Aging, USA)
cDNA Library Preparation: Yulan Piao and Minoru Ko (National Institute on Aging, NIH: <http://lgsun.grc.nia.nih.gov/cDNA/>)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAC Plate: 111 Row: e Column: 24
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753517.

Location/Qualifiers
1..2444
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x 129/Sv-CP"
/db_xref="taxon:10090"
/clone="MGC:60613 IMAGE:30008106"
/tissue_type="ES Cell, undifferentiated, mouse, R1 Embryonic stem cells"
/clone_lib="NIA Mouse Undifferentiated ES Cell cDNA Library (Long)"
/lab_host="DH10B"
/note="Vector: pSPORT1"
1..2444
/gene="Igf2bp1"
/note="synonyms: IMP-1, Neilsen, CRD-BP"
/db_xref="LocusID:140486"
/db_xref="MGI:1890357"
295..2028
/codon_start=1
/product="insulin-like growth factor 2, binding protein 1"
/protein_id="AAH51679.1"
/db_xref="GI:30354044"
/db_xref="LocusID:140486"
/translation="MNKLYIGNLNESTVPADLEKVFARHKISYSGQFLKSGYAFVDC PDEHWAKAIETFSKVLQKRLIEHSVPKORSKRIQIRNPQQRWEVLDSLLA QYGTVEQVNTSESTAVAVNTYSNREQTQATKLNHOLENHALKXVYIPDEQIT QCPNRRGGSGSGQPGQSGFVAGAPAKQPDVPIRLVPTQYVAGATGKCATI RNIKTQSTKIDVHRKNAGAAKASIVHSTPEGSSACKMILEIMHKEADTKTAG VPLKTAHNFTVGLIGKEGNLKVEQDTETKTITISLQDLTLNPTPTPVGAIE NCCRAEQIMKKVREAYENDVAASLQSHLIPGLNLAAGFPASSSSVPPPPVTG AAYSSFMQAPEQEMVQVFIIPAQAVGAIIGKGQHIKLSRFSASIKIAPPETPSK

gene
CDS

365 AGGAGTCGGAATAATACAGATCCGCAATATTCCACCTCAGCTCCGATGGGAAGTCTAGAT 424
168 AGCCTGCTGGTCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACACGAGAGTGAG 227
425 AGCCTGCTGGTCTCAGTATGTTACAGTAGAGAACTGTGAGCAAGTGAACACACTGAAAGTGAG 484
228 ACGGAGTGTGAATGTCACCTATTCCAAACGGGAGAGAGACACAGCAAGCCATCATGAAG 287
485 ACAGCGTGTGTCAACGTCACCTACTTAACCGGGAGCAGACACAGCAAGCTATCATGAAG 544
288 CTGAATGGCCACAGTGTGGAGAACATGCTCCCTGAAGGTCTCTACATCCCGATGAGCAG 347
545 CTAAATGGCCATCACTGAGAGAACATGCTCCCTGAAGGTCTCTACATCACTGATGAGCAG 604
348 ATAGCAGAGGACCTGAGAAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGC 407
605 ATACACAAGTCTGAGAAATGGGCGTCTGGAGGCTTTGGGTCTCGGGGCGAGCCCGG 664
408 CAGGGCTCACTGTGGCAGCGGGGGCCCCAGCCAGCAGCAGCAGTGGGATCGCCCTT 467
665 CAAGGCTCGCCCGTGGCAGCAGGGGCTCCAGCAAGCAGCAGCAGTGGGATCGCCCTT 724
468 CGGCTCCTGTGGTCCACCCAGTATGTGGTGCCTATTATTGCAAGGAGGGGCCACCATC 527
725 CGGCTCCTGTGGTCCACCCAGTATGTGGTGCCTATTATTGCAAGGAGGGGTGCCACCATC 784
528 CGCAACATCACAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGAGAAACGACGT 587
785 CGAAACATCACAAACAGACCCAGTCCAAATAGACGTGCATAGGAAGAGAAATCGCGGC 844
588 GCAGCTGAAAGCCATCAGTGTGCACTCCACCCCTGAGGGTGTCTCTCGCTTTGTAAG 647
845 GCTCGGAGAAAGCCATCAGCGTGCATTTCAACCCCTGAAAGGTGTCTCTCGCGCTGCAAG 904
648 ATGATCTTGGAGATTATCATTAAGAGGCTAAGGACACCAAAACGGCTGACAGGTTCCC 707
905 ATGATCTTGGAGATTATGCAAAAGAGGCAAGGACACCAAAACGGGAGATGAAGTTCCC 964
708 CTGAGATCTCTGGGCCATAATAACTTTGTAGGGCTCTCATTTGGCAAGGAGGACGGAAC 767
965 CTGAGATCTCTGGGTCAATAACATTTCTCGGGGACTCATTTGCAAGGAGGCGGAAC 1024
768 CTGAGAAAGGTAGAGCAAGATACCGAGACAAATATCACATCTCTCTGTTGCAAGCATT 827
1025 CTGAGAAAGGTGAGCAGGACACACAGACGAGATCACCATCTCATCTGCTCCAGGACCTC 1084
828 ACCCTTTACACCTTGAGGACCATCACTGTGAAGGGGCCATCGAGAAATTTGTTGCAAG 887
1085 ACGCTCTATAACCTTGAGGACCATCACTGTGAAGGGGCCATTTGAGAACTGTTGCAAG 1144
888 GCCGAGCAGGAAATAATGAAGAAAGTTCGGAGGGCTATGAGAAATGATGTGGCTGCCATG 947
1145 GCCGAGCAGGAGATCATGAAGAAAGTTCGAGAGGCTTACGAGAACGAGCTGGCGCCATG 1204
948 AGC-----TCTACCTGATCTGGCTGAAACCTGGCTGTGTAGGTCTTTTCCAGCT 1001
1205 AGCTTGAGTCCCACTCATCCCTGGCTTAACTGGCTGTGTAGGTCTCTTCCAGCT 1264
1002 TCATCCAGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTCTCCCTATAGTCC 1061
1265 TCATCCAGCGTGTCCCTCTCCAGAGTGTCACTGGGGCTGTCCCTATAGTCC 1324
1062 TTTATGAGGCTCCCGAGCAGGATGTTGCAAGTGTGTTATATCCCGCCAGGAGTGGC 1121
1325 TTATGAGGCTCCCGAGCAGGATGTTGCAAGTGTGTTATATCCCGCCAGGAGTGGC 1384
1122 GCCATCATCGGAAGAGGGGAGCAGATCAAGAGCTCTCCCGTTGGCCAGCGCTCC 1181
1385 GCCATCATCGGAAGAGGGGAGCAGATCAAGAGCTCTCCCGTTGGCCAGCGCTCC 1444
1182 ATCAAGATTGACCAACCGAAGACCTGACTCCAAAGTTTGTATGTTATCATCTGGA 1241
1445 ATCAAGATTGCTCCACCAAGAAACACCTGACTCCAAAGTTTGTATGTTGCTCATCTGGA 1504

Qy 1242 CCGCCAGAGGGCCCAATTCAAGGCTCAGGGAAGAAATCTATGCAAACTCAAGGAGAGAAC 1301
Db |||||
1505 CCCCAGAGGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGCAAACTAAAAGAGAGAA 1564
Qy 1302 TTCTTTGGTCCCAAGGAGGAGTGAAGCTGAGAGCCCAATACATGCTGTGCCAGCATCAGCA 1361
Db |||||
1565 TTCTTTGGTCCCAAGGAGGAGTGAAGCTAGAGACCCCAATACATGCGGTTCCGGCTTCAGCA 1624
Qy 1362 GCTGGCCGGGTCTATTGGCAAAAGTGGAAAAACGGTGAACGGTGTGCAAAATTTGACGCA 1421
Db |||||
1625 GCCGGCCGGTCTATCGGCAAAAGCGGCAAAACGGTGAATGAGTGCAGAACTTGACTGCA 1684
Qy 1422 GCTGAGGTGTGATACCAAGAGACAGACCCCTGATGAGAAGCAGCAGGTCTATCTGAAA 1481
Db |||||
1685 GCTGAGGTGTGATGCGCAAGAGACAGACCCCGGATGAGAAGCAGCAAGTCTATTGTTAAG 1744
Qy 1482 ATCATCCGACATTTCTATGCCAGTCAGATGCTCAACGGAAGATCCGAGACATCTCGGC 1541
Db |||||
1745 ATCATCGGACATTTCTATGCCAGCAGATGCTCAGCGAAGATCCGAGACATCTGGCT 1804
Qy 1542 CAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAAACAGGGCCAGGACGAGGAAAGTGA 1601
Db |||||
1805 CAAGTTAAGCAACAGCACCAAGGACAGAGCAACCTGGCCAGGCAACGAGGAAAGTGA 1864
Qy 1602 -CCAGCCCTCCCTGTCTCCCTTNGAGTCCAGGACACAAACGGGCGAGAA----- 1647
Db |||||
1865 CCGCCGCCCTCTCTGCCATTGGCTCCAGATCAGCAGGAGGAAACAGAACTCGAGGG 1924
Qy 1648 -----ATCGAGAGTGTCTCTCCCGGAGGCTCAGAAATGAGTGGAAATCCGGGAC 1699
Db |||||
1925 CGGGTGAAGGCGCGTGTGTTTCCAGCAGGCTCAGAAATGAGTGGAAATCCGGGAC 1983
Qy 1700 ACNTGGCGCGGCTGTAGATCAGGTTTCCCACTTGTATGAGAAAGATGTTCCAGTGAGG 1759
Db |||||
1984 ATTTGGGCTGCTGGAGATCAGGTTTGCACACTGATTTGAGAACAAATGTTCCAGTGAGG 2043
Qy 1760 AACCTTGATCTNTAGCCCCCAACACCCCAATGTCGCCCAACACTGNTGCCCTCGG 1819
Db |||||
2044 AATCTGATCTCTCGCCCCCAA--TTGAGCCAGCTGGCCACAGCCCAACCCCTTGGAAAT 2101
Qy 1820 GGTGTCAGAAATTTAGCGCAAGGCACTTTAAACGTGGATGTTTAAAGAGCTCTCCA 1879
Db |||||
2102 CACATTTCAATCATAGCTTGGTGTCTTTAAAGCTGGATGCT--TGAAGTTCTCA 2159
Qy 1880 GGGCCCCAACAAAGAGGTTGATCACACCTCAGTGGGAAGAAATAAATTTCTTCAGGT 1939
Db |||||
2160 GCCTCCATGGAAGGATGGTCCAGATCCAGTGGGGAAGAGAAATAAATTTCTTCAGGT 2219
Qy 1940 TTTA 1943
Db |||||
2220 TTTA 2223

RESULT 12
AF061569
LOCUS
DEFINITION Mus musculus coding region determinant binding protein mRNA,
complete cds.
ACCESSION AF061569
VERSION AF061569.1 GI:3273748
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2223)
AUTHORS Bernstein,P.L., Herrick,D.J., Prokipcak,R.D. and Ross,J.
TITLE Control of c-myc mRNA half-life in vitro by a protein capable of
binding to a coding region stability determinant
JOURNAL Genes Dev. 6 (4), 642-654 (1992)
MEDLINE 92217743
PUBMED 1559612

Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center, cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).

FEATURES

source
1. .2780
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3000789"
/cell_line="NT2"
/cell_type="Teratocarcinoma"
/clone_lib="NT2RP3"
/note="Cloning vector: pME18SPL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction"

ORIGIN

Query Match 71.4%; Score 1389; DB 9; Length 2780;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1405; Conservative 0; Mismatches 5; Indels 6; Gaps 1;
537 ACAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGAGAACCGAGGTGAGCTGAA 596
Db 1 ACAAACAGACCCAGTCCAAAGATAGACGTGCATAGGAAGAGAACCGAGGTGAGCTGAA 60
597 AAAGCCATCAGTGTGACTCCACCCCTGAGGCTGTCTCTCCGCTTGTAAAGATGATCTTG 656
Db 61 AAAGCCATCAGTGTGACTCCACCCCTGAGGCTGTCTCTCCGCTTGTAAAGATGATCTTG 120
657 GAGATTATGCATAAAGAGGCTAAGGACCAACAAACGGCTGACGAGGTTCCCTCTGAAGATC 716
Db 121 GAGATTATGCATAAAGAGGCTAAGGACCAACAAACGGCTGACGAGGTTCCCTCTGAAGATC 180
717 CTGGCCCATTAATAACTTTGTAGGCGGTCTCATTTGGCAAGGAAGACGGAACCTGAAGAAG 776
Db 181 CTGGCCCATTAATAACTTTGTAGGCGGTCTCATTTGGCAAGGAAGACGGAACCTGAAGAAG 240
777 GTAGAGCMAGATACCGAGACAAATAACCATCTCTCTCGTTGCAAGACCTTACCCCTTAC 836
Db 241 GTAGAGCAAGATACCGAGACAAATAACCATCTCTCTCGTTGCAAGACCTTACCCCTTAC 300
837 AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTTGTTGACGGGCGGAGCAG 896
Db 301 AACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAATTTGTTGACGGGCGGAGCAG 360
897 GAAATTAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGCTGCTGCATGAGC----- 950
Db 361 GAAATTAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGCTGCTGCATGAGC----- 420
951 TCTCAGCTGATCCCTGGCCCTGAACCTGGCTGTAGTGTCTTTCCAGCTTCCATCCAGC 1010
Db 421 TCTCAGCTGATCCCTGGCCCTGAACCTGGCTGTAGTGTCTTTCCAGCTTCCATCCAGC 480
1011 GCAGTCCCGCGGCTCCACGACAGGTTACTGGGCTGCTCCCTATAGTCTCTTTATGCGAG 1070
Db 481 GCAGTCCCGCGGCTCCACGACAGGTTACTGGGCTGCTCCCTATAGTCTCTTTATGCGAG 540
1071 GCTCCCGAGCAGAGATGTTGCGAGGTTTATCCCGCCCGAGGAGTGGGCGGCATCATC 1130
Db 541 GCTCCCGAGCAGAGATGTTGCGAGGTTTATCCCGCCCGAGGAGTGGGCGGCATCATC 600
1131 GGCAAGAGGGGCGAGACATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATT 1190
Db 601 GGCAAGAGGGGCGAGACATCAACAGCTCTCCCGTTTCCAGCGCTCCATCAAGATT 660
1191 GCACCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCTGACGACGCCAGAG 1250
Db 661 GCACCAACCCGAAACACCTGACTCCAAAGTTTCGTATGTTATCATCTGACGACGCCAGAG 720
1251 GCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTTCTTTGGT 1310
Db 721 GCCCAATTCAAGGCTCAGGGAAGATCTATGGCAAACTCAAGAGGAGAACTTCTTTGGT 780

1311 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCG 1370
Db 781 CCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGCGCG 840
1371 GTCATTGGCAAAAGGTGCAAAAACGGTGAACGAGTTTCAGAAATTTGACGGCAGCTGAGGTG 1430
Db 841 GTCATTGGCAAAAGGTGCAAAAACGGTGAACGAGTTTCAGAAATTTGACGGCAGCTGAGGTG 900
1431 GTAGTACCAAGAGACACAGACCCCTGTGATGAGAACGACGACGAGTTCATCTGAAAAATCATCGGA 1490
Db 901 GTAGTACCAAGAGACACAGACCCCTGTGATGAGAACGACGAGTTCATCTGAAAAATCATCGGA 960
1491 CATTTCATGACAGTCAGATGAGTCAACGGAAGATCCGAGACATCTCTGCCCCAGGTTAAG 1550
Db 961 CATTTCATGACAGTCAGATGAGTCAACGGAAGATCCGAGACATCTCTGCCCCAGGTTAAG 1020
1551 CAGCAGCATCAGAAAGGACAGAGTAACACAGGCCACGAGGACGAGGAAGTGAACACGACCCCT 1610
Db 1021 CAGCAGCATCAGAAAGGACAGAGTAACACAGGCCACGAGGACGAGGAAGTGAACACGACCCCT 1080
1611 CCTGTCCCTTNGAGTCCAGGACAAACACGCGGACGAAATCGAGAGTGTCTCTCCCGGC 1670
Db 1081 CCTGTCCCTTNGAGTCCAGGACAAACACGCGGACGAAATCGAGAGTGTCTCTCCCGGC 1140
1671 AGSCCTGGAATCAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGCCC 1730
Db 1141 AGSCCTGGAATCAGTGGGAATCCGGGACACNTGGCGGGCTGTAGATCAGGTTTGCCC 1200
1731 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTNTCAGCCCCAAACACCCACC 1790
Db 1201 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCTGATCTNTCAGCCCCAAACACCCACC 1260
1791 CAATTGGCCCCAACACTGTCTGCGGCTCGGGGTGTAGAAATTTCTAGCAAGGCACTTTT 1850
Db 1261 CAATTGGCCCCAACACTGTCTGCGGCTCGGGGTGTAGAAATTTCTAGCAAGGCACTTTT 1320
1851 AAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCACGAGAGGTTGGATCACACCTCAG 1910
Db 1321 AAACGTGGATTGTTTAAAGAGCTCTCCAGGCCCCACGAGAGGTTGGATCACACCTCAG 1380
1911 TGGGAAGAAAAATATAAATTTCTTTCAGGTTTAAAA 1946
Db 1381 TGGGAAGAAAAATATAAATTTCTTTCAGGTTTAAAA 1416
RESULT 11
ARI60244
LOCUS 2224 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6255055.
ACCESSION ARI60244
VERSION ARI60244.1 GI:16223887
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2224)
AUTHORS Ross,J.
TITLE c-myc coding region determinant-binding protein (CRD-BP) and its
nucleic acid sequence
JOURNAL Patent: US 6255055-A 1 03-JUL-2001;
FEATURES Location/Qualifiers
source 1. .2224
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 66.9%; Score 1301; DB 6; Length 2224;
Best Local Similarity 83.5%; Pred. No. 1.9e-308;
Matches 1556; Conservative 0; Mismatches 275; Indels 33; Gaps 6;
108 AGGAGCGGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAAGTACTGGAC 167

PI	TETSUJI OTSUKI,HISASHI KOGA
PC	C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,	10,
PC,	C12P21/02,C12Q1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer	for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers	
FT	CDS (112)..(1068).
FEATURES	
source	1..2780
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
ORIGIN	
Query Match	71.4%; Score 1389; DB 6; Length 2780;
Best Local Similarity	99.2%; Pred. No. 0;
Matches 1405; Conservative	0; Mismatches 5; Indels 6; Gaps 1;
Qy	537 ACAAACAGACCCAGTCCAGATAGACGTGATAGGAAGGAGAACGCGAGTGCAGTGAA 596
Db	1 ACAAACAGACCCAGTCCAGATAGACGTGATAGGAAGGAGAACGCGAGTGCAGTGAA 60
Qy	597 AAAGCCATCAGTGTGCATCTCACCCCTGAGGGCTGCTCCGCTTGTGAAGATGATCTTG 656
Db	61 AAAGCCATCAGTGTGCATCTCACCCCTGAGGGCTGCTCCGCTTGTGAAGATGATCTTG 120
Qy	657 GAGATTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGGTTCCCTCGAAGATC 716
Db	121 GAGATTATGCATAAAGAGGCTTAAGACACCAAAACGGCTGACGAGTTCCTCGAAGATC 180
Qy	717 CTGGCCCAATAAATCTTTGTAGGGCGTCTCATTTGCAAGGAGGACGGAACCTGAAGAAG 776
Db	181 CTGGCCCAATAAATCTTTGTAGGGCGTCTCATTTGCAAGGAGGACGGAACCTGAAGAAG 240
Qy	777 GTAGAGCAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTTACCTTTTAC 836
Db	241 GTAGAGCAGATACCGAGACAAAATCACCATCTCTCTGTTGCAAGACCTTACCTTTTAC 300
Qy	837 AACCTGAGAGGACCATCAGTGTGAAGGGGCCATCGAGAAATGTTGCAAGGCCCGAGCAG 896
Db	301 AACCTGAGAGGACCATCAGTGTGAAGGGGCCATCGAGAAATGTTGCAAGGCCCGAGCAG 360
Qy	897 GAAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGGCTGCCATGAGC----- 950
Db	361 GAAATAATGAAGAAAGTTCCGGAGGCCCTATGAGAAATGATGGCTGCCATGAGCCTGCAG 420
Qy	951 TCTCACCTGATCCCTGGCTGAACTGCTGCTGATAGTCTTTTCCGAGCTTCATCCAGC 1010
Db	421 TCTCACCTGATCCCTGGCTGAACTGCTGCTGATAGTCTTTTCCGAGCTTCATCCAGC 480
Qy	1011 GCAGTCCCGCGCTCCAGCAGCGCTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAG 1070
Db	481 GCAGTCCCGCGCTCCAGCAGCGCTTACTGGGGCTGCTCCCTATAGCTCCTTTATGCAG 540
Qy	1071 GCTCCCGAGCAGGAGTGTGCAGTGTATTATCCCGCCAGGCGAGTGGCGCCATCATC 1130
Db	541 GCTCCCGAGCAGGAGTGTGCAGTGTATTATCCCGCCAGGCGAGTGGCGCCATCATC 600
Qy	1131 GCGAAGAGGGGCGAGCAGATCAACAGCTCTCCCGTTTGCAGCGCTCCATCAAGATT 1190
Db	601 GCGAAGAGGGGCGAGCAGATCAACAGCTCTCCCGTTTGCAGCGCTCCATCAAGATT 660
Qy	1191 GCACCACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCACTGGACCGCCAGAG 1250
Db	661 GCACCACCCGAAACACCTGACTCCAAAGTTGATGTTATCATCACTGGACCGCCAGAG 720
Qy	1251 GCCCAATTCAAGGCTCAGGGAAGATCTATGCAAACTCAAGAGGAGAACTTCTTTGGT 1310
Db	721 GCCCAATTCAAGGCTCAGGGAAGATCTATGCAAACTCAAGAGGAGAACTTCTTTGGT 780
Qy	1311 CCCAAGGAGGAGTCAAGCTGGAGACCCACATACGTGTGCCAGATCAGCAGCTGGCCGG 1370

Db	781 CCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCCGG 840
Qy	1371 GTCATTGGCAAGGTGGAATAACCGGTGAACGAGTGTGAGAAATTTGACCGGAGCTGAGGTG 1430
Db	841 GTCATTGGCAAGGTGGAATAACCGGTGAACGAGTGTGAGAAATTTGACCGGAGCTGAGGTG 900
Qy	1431 GTAGTACCAAGAGACGAGACCCCTGATGAGAACGACGAGGTCACTGTAATAATCATCGGA 1490
Db	901 GTAGTACCAAGAGACGAGACCCCTGATGAGAACGACGAGGTCACTGTAATAATCATCGGA 960
Qy	1491 CATTTCTATGCGAGTTCAGATGGCTCAACGGAAGATCCGAGACATCCTTGGCCCGAGTTAAG 1550
Db	961 CATTTCTATGCGAGTTCAGATGGCTCAACGGAAGATCCGAGACATCCTTGGCCCGAGTTAAG 1020
Qy	1551 CAGCAGCATCAGAAAGGACAGAGTAACGAGCCCGCAGGACGAGGAAGTGCACGCCCT 1610
Db	1021 CAGCAGCATCAGAAAGGACAGAGTAACGAGCCCGCAGGACGAGGAAGTGCACGCCCT 1080
Qy	1611 CCCTGTCCCTTTNGAGTCCAGGACACACGCGGCGAATAATCGAGAGTGTCTCTCCCGGCG 1670
Db	1081 CCCTGTCCCTTTNGAGTCCAGGACACACGCGGCGAATAATCGAGAGTGTCTCTCCCGGCG 1140
Qy	1671 AGGCTGAGAAATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGCC 1730
Db	1141 AGGCTGAGAAATGAGTGGGAATCCGGGACACCTGGGCGGGCTGTAGATCAGGTTTGCC 1200
Qy	1731 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTTNTCAGCCCCAAACACCCACC 1790
Db	1201 ACTTGATTGAGAAAGATGTTCCAGTGAGGAACCCCTGATCTTNTCAGCCCCAAACACCCACC 1260
Qy	1791 CAATTGGCCCAACACTGTTNTGCCCTCGGGGTGTCAGAAATTTAGCCGACAGGCACTTTT 1850
Db	1261 CAATTGGCCCAACACTGTTNTGCCCTCGGGGTGTCAGAAATTTAGCCGACAGGCACTTTT 1320
Qy	1851 AAACCTGTGATTGTTTAAAGAAAGCTCTCCAGGCCCCCACCAGAGGGTGGATCACACTCAG 1910
Db	1321 AAACCTGTGATTGTTTAAAGAAAGCTCTCCAGGCCCCCACCAGAGGGTGGATCACACTCAG 1380
Qy	1911 TGGGAAGAAAAATAAAATTTCTTCAAGTTTTTAAAA 1946
Db	1381 TGGGAAGAAAAATAAAATTTCTTCAAGTTTTTAAAA 1416
RESULT 10	
AK074915	2780 bp mRNA linear PRI 03-SEP-2002
LOCUS	Homo sapiens cDNA FLJ90434 fis, clone NT2RP3000789, highly similar
DEFINITION	to Homo sapiens IGF-II mRNA-binding protein 1 (IMP-1) mRNA.
ACCESSION	AK074915
VERSION	AK074915.1 GI:22760672
KEYWORDS	oligo capping; fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	1
	Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
	Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
	Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
	Kojima,S., Nagahari,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
	Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
	Ninomiya,K.
	NEDO human cDNA sequencing project
TITLE	
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2780)
AUTHORS	Isogai,T. and Otsuki,T.
TITLE	Direct Submission
JOURNAL	Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
	Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
	(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
	NEDO human cDNA sequencing project supported by Ministry of
	Economy, Trade and Industry of Japan; cDNA full insert sequencing:
COMMENT	Research Association for Biotechnology; cDNA library construction:

/db_xref="taxon:9606"									
ORIGIN									
Query Match 85.8%; Score 1669.2; DB 6; Length 1708;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
QY	268	CCAGGCAAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCGCTGAAGGTCT	327						
Db	30	CCCGGGAGCCATCATGAAGCTGAATGGCCACCAGTTGGAGAACCATGCGCTGAAGGTCT	89						
QY	328	CCTACATCCCCGATGACGATAGCACAGGACCTGAGAAATGGCGCGAGGGGGCTTG	387						
Db	90	CCTACATCCCCGATGACGATAGCACAGGACCTGAGAAATGGCGCGAGGGGGCTTG	149						
QY	388	GCTCTCGGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGACCAAGCAGC	447						
Db	150	GCTCTCGGGTCAGCCCGCCAGGGCTCACCTGTGGCAGCGGGGGCCCCAGACCAAGCAGC	209						
QY	448	AGCAAGTGGACATCCCGCTTCGGCTCCTGTGTGCCACCCAGTATGTGGGTGCCATTATG	507						
Db	210	AGCAAGTGGACATCCCGCTTCGGCTCCTGTGTGCCACCCAGTATGTGGGTGCCATTATG	269						
QY	508	GCAGAGGGGGCCACCATCCGCAACATCAACAAAGAGCCAGTCCAGATAGAGGTGC	567						
Db	270	GCAGAGGGGGGGCCACCATCCGCAACATCAACAAAGAGCCAGTCCAGTCCAGATAGAGGTGC	329						
QY	568	ATAGGAAGAGAAACGAGGTGCAGCTGAAAGAGCCATCAGTGTGCATCCACCCCTGAGG	627						
Db	330	ATAGGAAGAGAAACGAGGTGCAGCTGAAAGAGCCATCAGTGTGCATCCACCCCTGAGG	389						
QY	628	GCTGCTCCTCCGTTGTAAGATGATCTTGGAGATATTGCAATAAGAGGCTTAAGACCA	687						
Db	390	GCTGCTCCTCCGTTGTAAGATGATCTTGGAGATATTGCAATAAGAGGCTTAAGACCA	449						
QY	688	AAACGGCTGACAGGTTCCCTGAGATCCTGGCCCATATATCTTTGTAGGGCGTCTCA	747						
Db	450	AAACGGCTGACAGGTTTCCCTGAGATCCTGGCCCATATATATCTTTGTAGGGCGTCTCA	509						
QY	748	TTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCA	807						
Db	510	TTGGCAAGAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAAAATCACCA	569						
QY	808	TCCTCTGTTGGAAGACTTACCTTTACAACTTGAAGAGACATCATTCTGTAAGGGGG	867						
Db	570	TCCTCTGTTGGAAGACTTACCTTTTACAACTTGAAGAGACCATCATTCTGTAAGGGGG	629						
QY	868	CCATCGAAGATTGTTCAGGGCCGAGCAGGAAATATGAAGAAAGTTTCGGGAGGCCCTATG	927						
Db	630	CCATCGAAGATTGTTCAGGGCCGAGCAGGAAATATGAAGAAAGTTTCGGGAGGCCCTATG	689						
QY	928	AGAATGATGTGGTGCATGAGCTCTCACCTGATCCCTGGCCCTGAACCTGGCTGTAG	987						
Db	690	AGAATGATGTGGTGCATGAGCTCTCACCTGATCCCTGGCCCTGAACCTGGCTGTAG	749						
QY	988	GTCTTTTCCAGCTTCTACAGCGCAGTCCCGCGCCTCCAGCAGGCTTACTGGGGCTG	1047						
Db	750	GTCTTTTCCAGCTTCTATCCAGCGCAGTCCCGCGCCTCCAGCAGGCTTACTGGGGCTG	809						
QY	1048	CTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTATCCCG	1107						
Db	810	CTCCCTATAGCTCTTTATGAGGCTCCCGAGCAGGAGATGGTGCAGGTGTTATCCCG	869						
QY	1108	CCAGGAGTGGGCGCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGT	1167						
Db	870	CCAGGAGTGGGCGCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGT	929						
QY	1168	TTGCGAGGCTTCATCAAGATTGCACACCCGAAACACCTGATCCAAAGTTCGTATGG	1227						
Db	930	TTGCGAGGCTTCATCAAGATTGCACACCCGAAACACCTGATCCAAAGTTCGTATGG	989						
QY	1228	TTATCATCACTGGACCGCCAGGCCCCCAATTCAGGCTCAGGGAAGATCTATGGCAAC	1287						

Db	990	TTATCATCACTGGACCGCCAGAGGCCCAATTCAGAGCTCAGGGAAGAAATCTATGGCAAC	1049
QY	1288	TCAAGGAGGAGAACTTTCTTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTG	1347
Db	1050	TCAAGGAGGAGAACTTTCTTTTGGTCCCAAGGAGAGTGAAGCTGGAGACCCACATACGTG	1109
QY	1348	TGCCAGATCAGCAGCTGCGCGGTCAITGGCAAGGTGGAAAACCGTGAACAGTTGC	1407
Db	1110	TGCCAGATCAGCAGCTGCGCGGTCAITGGCAAGGTGGAAAACCGTGAACAGTTGC	1169
QY	1408	AGAAATTTGACGCGCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACC	1467
Db	1170	AGAAATTTGACGCGCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACC	1229
QY	1468	AGTCTCATCTGAAAATCATCGGACATTTCTATGCCAGTCCAGATGGCTCAACGGAAGATCC	1527
Db	1230	AGTCTCATCTGAAAATCATCGGACATTTCTATGCCAGTCCAGATGGCTCAACGGAAGATCC	1289
QY	1528	GAGACATCTGGCCAGGTTAAGCAGCAGATCAGAAAGGACAGATTAACGAGGCCAGG	1587
Db	1290	GAGACATCTGGCCAGGTTAAGCAGCAGATCAGAAAGGACAGATTAACGAGGCCAGG	1349
QY	1588	CACGAGGAGAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGCGGAGAA	1647
Db	1350	CACGAGGAGAGTGACCCAGCCCTCCCTGTCCCTTNGAGTCCAGGACAAACACGCGGAGAA	1409
QY	1648	ATCGAGAGTGTCTCTCCCGGAGCCCTGAGAAATGAGTGGGAATCCGGGACACNTGGGC	1707
Db	1410	ATCGAGAGTGTCTCTCCCGGAGCCCTGAGAAATGAGTGGGAATCCGGGACACNTGGGC	1469
QY	1708	CGGGCTGTAGATCAGGTTTGGCCACTTGTGAAAGATGTTCAGTCCAGTCCAGTCCAGT	1767
Db	1470	CGGGCTGTAGATCAGGTTTGGCCACTTGTGAAAGATGTTCAGTCCAGTCCAGTCCAGT	1529
QY	1768	TCNTCAGCCCAACACCCCAATTCGCCCAACACTGTNTGCCCTCGGGGTGTCTAG	1827
Db	1530	TCNTCAGCCCAACACCCCAATTCGCCCAACACTGTNTGCCCTCGGGGTGTCTAG	1589
QY	1828	AAATTNTAGCGAAGGACATTTTAAACGTGGATGTTTAAAGAGCTTCCAGGCCCCAC	1887
Db	1590	AAATTNTAGCGAAGGACATTTTAAACGTGGATGTTTAAAGAGCTTCCAGGCCCCAC	1649
QY	1888	CAGAGGCTGTATCAGACCTCAGTGGGAGAAATAAATTTCTTCCAGTTTAAAA	1946
Db	1650	CAGAGGCTGTATCAGACCTCAGTGGGAGAAATAAATTTCTTCCAGTTTAAAA	1708

RESULT 9	BD127811	2780 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	Primer for synthesizing full-length cDNA and use thereof.				
DEFINITION	BD127811				
ACCESSION	BD127811.1 GI:23222756				
VERSION	JP 2002017375-A/3242				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2780)				
AUTHORS	Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and Koga, H.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002017375-A 3242 22-JAN-2002;				
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) EN JP 2002017375-A/3242 PD 22-JAN-2002 PF 07-JUL-2000 JP 2000253172 PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO ISHII, PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI SHINICHI KOJIMA,				

388 GCTCTCGGGGTACGCCCGCCAGGGCTCACTGTGGGAGCGGGGGCCCCAGCGAAGCAGC 447
Db GCTCTCGGGGTACGCCCGCCAGGGCTCACTGTGGGAGCGGGGGCCCCAGCGAAGCAGC 209
448 AGCAAGTGGACATCCCGCTTGGGCTCTGTGGTCCCGCCAGCCAGTATGTGGTCCCAATTATG 507
Db AGCAAGTGGACATCCCGCTTGGGCTCTGTGGTCCCGCCAGCCAGTATGTGGTCCCAATTATG 269
508 GCAAGGAGGGGGCCACCATCCGCAACATCAAAACAGACCAGCCAGTCCAAAGATAGACGTGC 567
Db GCAAGGAGGGGGCCACCATCCGCAACATCAAAACAGACCAGCCAGTCCAAAGATAGACGTGC 329
568 ATAGGAAGGAGAAACGAGGTGAGCTGAAAAGGCAATCAGTGTGCACTCCACCCCTGAGG 627
Db ATAGGAAGGAGAAACGAGGTGAGCTGAAAAGGCAATCAGTGTGCACTCCACCCCTGAGG 389
628 GCTCCTCTCGCTGTGAAGATGATCTTGGAGATTTATGCAATAAGAGGCTAAGGACACCA 687
Db GCTCCTCTCGCTGTGAAGATGATCTTGGAGATTTATGCAATAAGAGGCTAAGGACACCA 449
688 AAACGGGTGACGAGGTTCCTTGAAGATCCTTGGGCCATAATAACTTTGTAGGGGTCTCA 747
Db AAACGGGTGACGAGGTTCCTTGAAGATCCTTGGGCCATAATAACTTTGTAGGGGTCTCA 509
748 TTGCAAGGAAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAATACCA 807
Db TTGCAAGGAAGGACGGAACCTGAAGAGGTAGAGCAAGATACCGAGACAAATACCA 569
808 TCTCCTCTGTCAGACCTTACCTTTACACCTTGAGAGGACCATCACTGTGAAGGGG 867
Db TCTCCTCTGTCAGACCTTACCTTTACACCTTGAGAGGACCATCACTGTGAAGGGG 629
868 CCATCGAATTTGTTGAGGCGCCGAGCAGGAATAATAAGAAAGTTCCGGAGGCGCTATG 927
Db CCATCGAATTTGTTGAGGCGCCGAGCAGGAATAATAAGAAAGTTCCGGAGGCGCTATG 689
928 AGAATGATGTGGTCCGATGAGCTCTACCTGATCCCTGGCTGAACTTGGCTGTGTAG 987
Db AGAATGATGTGGTCCGATGAGCTCTACCTGATCCCTGGCTGAACTTGGCTGTGTAG 749
988 GTCTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGCGTTACTGGGGGTG 1047
Db GTCTTTCCAGCTTCATCCAGCGCAGTCCCGCGCTCCAGCAGCGTTACTGGGGGTG 809
1048 CTCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGATGAGGTGTTATTCCTCGG 1107
Db CTCCTATAGTCTCTTTATGAGGCTCCCGAGCAGGAGATGATGAGGTGTTATTCCTCGG 869
1108 CCCAGGAGTGGGCGCCATATCGGCAAGAGAGGGGAGCAGATCAACAGCTCTCCCGGT 1167
Db CCCAGGAGTGGGCGCCATATCGGCAAGAGAGGGGAGCAGATCAACAGCTCTCCCGGT 929
1168 TTGCCAGCGCTCCATCAAGATTCACACCCGCAACACCTGACTCCAAAGTTCGTATGG 1227
Db TTGCCAGCGCTCCATCAAGATTCACACCCGCAACACCTGACTCCAAAGTTCGTATGG 989
1228 TTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGATCTATGCAAAAC 1287
Db TTATCATCACTGGACCGCCAGAGGCCCAATTCAAGGCTCAGGGAAGATCTATGCAAAAC 1049
1288 TCAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG 1347
Db TCAGGAGGAGAACTTCTTTGGTCCCAAGGAGGAAGTGAAGCTGGAGACCCACATACGTG 1109
1348 TGCCAGCATCAGCAGCTGCGGGGTCAATTGGCAAGGTGGAATAACGGTGAACGAGTTGC 1407
Db TGCCAGCATCAGCAGCTGCGGGGTCAATTGGCAAGGTGGAATAACGGTGAACGAGTTGC 1169
1408 AGAATTTGACGGCAGCTGAGGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACC 1467
Db AGAATTTGACGGCAGCTGAGGTGTGTAGTACCAAGAGACAGACCCCTGATGAGAACGACC 1229
1468 AGGTCACTCGTGAATAATCATCGACATTTCTATGCCAGTCAGTGGCTCAACGGAAGATCC 1527

Db AGGTCACTCGTGAATAATCATCGACATTTCTATGCCAGTCAGATGGCTCAACGGAAGATCC 1289
QY GAGACATCTCTGGCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAAACAGGCCAGG 1587
Db GAGACATCTCTGGCCAGGTTAAGCAGCAGCATCAGAAGGACAGAGTAAACAGGCCAGG 1349
QY CACGAGGAAGTACCGAGCCCTCCTGTCTCTTNGAGTCCAGGACAAACAGCGGCAGAA 1647
Db CACGAGGAAGTACCGAGCCCTCCTGTCTCTTNGAGTCCAGGACAAACAGCGGCAGAA 1409
QY ATCCAGAGTGTCTCTCCCGCAGGCTGAGATAGTGGAAATCCGGGACACNTGGC 1707
Db ATCCAGAGTGTCTCTCCCGCAGGCTGAGATAGTGGAAATCCGGGACACNTGGC 1469
QY CGGCTGTAGATCAGGTTTGCCCACTTGATTTGAGAAAGATGTTCCAGTGAGGAACCTTGA 1767
Db CGGCTGTAGATCAGGTTTGCCCACTTGATTTGAGAAAGATGTTCCAGTGAGGAACCTTGA 1529
QY TCTNTCAGCCCCCAACACCCACCAATTTGGCCCACTGTNTGCCCTCGGGGTGTGAG 1827
Db TCTNTCAGCCCCCAACACCCACCAATTTGGCCCACTGTNTGCCCTCGGGGTGTGAG 1589
QY AAATNTAGCCCAAGGACATTTTAAACGTGATTTTAAAGAGCTCTCCAGGCCCCAC 1887
Db AAATNTAGCCCAAGGACATTTTAAACGTGATTTTAAAGAGCTCTCCAGGCCCCAC 1649
QY CAAGAGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTTTAAA 1946
Db CAAGAGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTTTAAA 1708

RESULT 8
BD209924 1708 bp DNA linear PAT 17-JUL-2003
LOCUS Isolated nucleic acid molecules encoding cancer-associated
DEFINITION antigens, these antigens and method of using the same.
ACCESSION BD209924.1 GI:33019694
VERSION BD209924.1
KEYWORDS JP 2002512049-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1708)
AUTHORS Chen, Y. T., Gure, A., Tsang, S., Stockert, E., Jager, E., Knuth, A. and
Old, L. J.
TITLE Isolated nucleic acid molecules encoding cancer-associated
antigens, these antigens and method of using the same
JOURNAL Patent: JP 2002512049-A 3 23-APR-2002;
COMMENT LUDWIG INSTITUTE FOR CANCER RESEARCH
OS Homo sapiens (human)
PN JP 2002512049-A/3
PD 23-APR-2002
PF 16-MAR-1999 JP 2000545030
PR 17-APR-1998 US 09/061709
PI YAO TSENG CHEN, ALI GURE, SOLAM TSANG, ELISABETH STOCKERT, ELKE
PI JAGER, PI
PI ALEXANDER KNUTH, LLOYD J OLD
PC C12N15/09, A61K35/12, A61K39/00, A61K39/39, A61P35/00, C07K16/32,
PC C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/53, G01N33/PC
574,
PC C12N15/00, C12N5/00
CC n-not determined(1384, 1464, 1533, 1571, 1595)
FH Key Location/Qualifiers
FT source 1..1708
FT /organism='Homo sapiens (human)'.
FT Location/Qualifiers
1..1708
/organism='Homo sapiens'
/mol_type='genomic DNA'


```
Db 30 CCGGGAGGACATCATGAAGCTGATGGCCACAGATTGGAGAACCATAGCCCTGAAGGTCT 89
QY 328 CTTACATCCCAGATGAGCAGATAGCA CAGGAGACCTGAGAAATGGCGCGCGAGGGGGCTTTG 387
Db 90 CTTACATCCCAGATGAGCAGATAGCA CAGGAGACCTGAGAAATGGCGCGCGAGGGGGCTTTG 149
QY 388 GCTCTCGGGGTGAGCCCGGCGAGGGCTCACTGTGGCAGAGGGGGCCCGAGCCAGAGCAGC 447
Db 150 GCTCTCGGGGTGAGCCCGGCGAGGGCTCACCTGTGGCAGAGGGGGCCCGAGCCAGAGCAGC 209
QY 448 AGCAAGTGAACATCCCCCTTCCGCTCTCTGTGTGGCCACCCAGTATGTGGGTGCCATTATTG 507
Db 210 AGCAAGTGAACATCCCCCTTCCGCTCTCTGTGTGGCCACCCAGTATGTGGGTGCCATTATTG 269
QY 508 GCAAGAGGGGGGACCAATCCGCAACATCA CAAACAGACCCAGTCCAAAGATAGACGTGC 567
Db 270 GCAAGAGGGGGGACCAATCCGCAACATCA CAAACAGACCCAGTCCAAAGATAGACGTGC 329
QY 568 ATAGGAGGAGGAGCGGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGG 627
Db 330 ATAGGAGGAGGAGCGGAGGTGCAGCTGAAAAGCCATCAGTGTGCATCTCCACCCCTGAGG 389
QY 628 GCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGCACCA 687
Db 390 GCTGCTCTCCGCTTGTAAAGATGATCTTGGAGATTATGCATAAAGAGGCTAAGCACCA 449
QY 688 AAGCGGTGAGGAGGTTCCTCGAAGATCTCGGCCAATAATACTTTGTAGGGCGTCTCA 747
Db 450 AAGCGGTGAGGAGGTTCCTCGAAGATCTCGGCCAATAATACTTTGTAGGGCGTCTCA 509
QY 748 TTGCGAAGGAGGACGGAACCTGAGAGGTAGAGCAAGATACCGAGACAAAATCACCA 807
Db 510 TTGCGAAGGAGGACGGAACCTGAGAGGTAGAGCAAGATACCGAGACAAAATCACCA 569
QY 808 TCTCCTCGTTCGAAGACCTTACCCCTTTACACCCCTGAGAGGACCATCAGTGTGAAGGGGG 867
Db 570 TCTCCTCGTTCGAAGACCTTACCCCTTTACACCCCTGAGAGGACCATCAGTGTGAAGGGGG 629
QY 868 CCAATGAGAAATTTGTCAGGGGCGAGCAGGAAATATGAAAGAAATTTGCGAGGCGCTATG 927
Db 630 CCAATGAGAAATTTGTCAGGGGCGAGCAGGAAATATGAAAGAAATTTGCGAGGCGCTATG 689
QY 928 AGAATGATGCGTGCATCAGCTCTCACCTGATCCCTGCGCTGAACTGGCTGCTGTAG 987
Db 690 AGAATGATGCGTGCATCAGCTCTCACCTGATCCCTGCGCTGAACTGGCTGCTGTAG 749
QY 988 GTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 1047
Db 750 GTCTTTTCCAGCTTCATCCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGGCTG 809
QY 1048 CTCCTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGTTGTCAGGTGTTATCCCG 1107
Db 810 CTCCTATAGCTCTTTATGAGGCTCCCGAGCAGAGATGTTGTCAGGTGTTATCCCG 869
QY 1108 CCAGGCGAGTGGCGGCTCATCCGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGT 1167
Db 870 CCAGGCGAGTGGCGGCTCATCCGCAAGAGGGGCGAGCATCAACAGCTCTCCCGGT 929
QY 1168 TTGCCAGGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTTCGTATGG 1227
Db 930 TTGCCAGGCGCTCCATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTTCGTATGG 989
QY 1228 TTATCATCTGACCGCCAGAGGCCCAATTTAAGGCTAGGGAAGATCTATGGCAAC 1287
Db 990 TTATCATCTGACCGCCAGAGGCCCAATTTAAGGCTAGGGAAGATCTATGGCAAC 1049
QY 1288 TCAAGAGGAGAACTCTTTTGTCCCAAGGAGGAGTGAAGCTGGAGCCCAATACGTG 1347
Db 1050 TCAAGAGGAGAACTCTTTTGTCCCAAGGAGGAGTGAAGCTGGAGCCCAATACGTG 1109
QY 1348 TGCCAGCATCAGCAGCTGGCGGGGTCACTTGGCAAAAGGTGAAAAACCGGTGAACGAGTTGC 1407
```

```
Db 1110 TGCCAGCATCAGCAGCTGGCCGGGTCACTTGGCAAAAGGTGGAATAACCGTGAAACGAGTTGC 1169
QY 1408 AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAAGACCCCTGATGAGAACGACC 1467
Db 1170 AGAATTTGACGGCAGCTGAGGTGGTAGTACCAAGAGACCAAGACCCCTGATGAGAACGACC 1229
QY 1468 AGTCACTGTAAGAAATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACCGAAGATCC 1527
Db 1230 AGTCACTGTAAGAAATCATCGGACATTTCTATGCCAGTCAAGTGGCTCAACCGAAGATCC 1289
QY 1528 GACACATCTCTGGCCAGGTAAAGCAGCAGCATCAGAAAGGACAGAGTAACACGAGCCGAGG 1587
Db 1290 GAGACATCTCTGGCCAGGTAAAGCAGCAGCATCAGAAAGGACAGAGTAACACGAGCCGAGG 1349
QY 1588 CACGAGGAGTGCACGAGCCCTCCCTGCTCCCTTNGAGTCCAGGACAAACAAACGGGAGAA 1647
Db 1350 CACGAGGAGTGCACGAGCCCTCCCTGCTCCCTTNGAGTCCAGGACAAACAAACGGGAGAA 1409
QY 1648 ATCAGAGTGTGCTCTCCCGGAGGCGCTGAGAAATGAGTGGGAATCGGGACACNTGGGC 1707
Db 1410 ATCAGAGTGTGCTCTCCCGGAGGCGCTGAGAAATGAGTGGGAATCGGGACACNTGGGC 1469
QY 1708 CGGCTGTAGATCAGGTTTGCCTTGTGAGAAAGATGTTCCAGTGAAGAACCTGTA 1767
Db 1470 CGGCTGTAGATCAGGTTTGCCTTGTGAGAAAGATGTTCCAGTGAAGAACCTGTA 1529
QY 1768 TCTNTCAGCCCAACACCAACCTTAAAGTGGCCCAACCTGNTGCCCTCGGGGTGTGAG 1827
Db 1530 TCTNTCAGCCCAACACCAACCTTAAAGTGGCCCAACCTGNTGCCCTCGGGGTGTGAG 1589
QY 1828 AATTTTAGGCAAGGACCTTTTAAAGTGGATGTTTAAAGAGTCTCCAGGCCCCAC 1887
Db 1590 AATTTTAGGCAAGGACCTTTTAAAGTGGATGTTTAAAGAGTCTCCAGGCCCCAC 1649
QY 1888 CAAGAGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTAAAA 1946
Db 1650 CAAGAGGTGGATCACACCTCAGTGGGAAGAAAATAAAATTTCTTCAGGTTTAAAA 1708
```

RESULT 7

```
AR343073 LOCUS 1708 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 5 from patent US 6576756.
ACCESSION AR343073
VERSION AR343073.1 GI:33738475
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1708)
AUTHORS Chen,Y.-F., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
JOURNAL Patent: US 6576756-A 5 10-JUN-2003;
FEATURES Location/Qualifiers
source 1..1708
/organism="unknown"
/mol_type="genomic DNA"
```

ORIGIN

```
Query Match 85.8%; Score 1669.2; DB 6; Length 1708;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 268 CCAGGCAAGCCATCATGAAGCTGAATGGCCACACAGTTGGAGAACCATGCCCTGAAGGTCT 327
Db 30 CCGGGGAGCCATCATGAAGCTGAATGGCCACACAGTTGGAGAACCATGCCCTGAAGGTCT 89
QY 328 CTTACATCCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGGCGCGAGGGGGCTTTG 387
Db 90 CTTACATCCCCGATGAGCAGATAGCACAGGACCTGAGAAATGGGCGCGAGGGGGCTTTG 149
```


Db 600 AGCTGCTGGCTCAGTATGTTAGTAGAGAACTGTGAGCAAGTGAACACCGAGGTGAG 659
 Qy 228 ACGGAGTGTGTAATGTCACTTATTCACACCGGGAGCAGACCGAGCAAGCCATCATGAAG 287
 Db 660 ACGGAGTGTGTAATGTCACTTATTCACACCGGGAGCAGACCGAGCAAGCCATCATGAAG 719
 Qy 288 CTGAATGGCCACCAAGTTGGAGAACATGCTCCCTGAAGGTCTCTACATCCCGGATGAGCAG 347
 Db 720 CTGAATGGCCACCAAGTTGGAGAACATGCTCCCTGAAGGTCTCTACATCCCGGATGAGCAG 779
 Qy 348 ATAGCAGAGGACCTGGAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCG 407
 Db 780 ATAGCAGAGGACCTGGAATGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCG 839
 Qy 408 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCCCTT 467
 Db 840 CAGGGCTCACCTGTGGCAGCGGGGCCCCAGCCAGCAGCAGCAAGTGGACATCCCCCTT 899
 Qy 468 CGGCTCTGGTGGCCACCCAGTATGTGGTGGCCATTATTGGCAAGGAGGGGCCACCATC 527
 Db 900 CGGCTCTGGTGGCCACCCAGTATGTGGTGGCCATTATTGGCAAGGAGGGGCCACCATC 959
 Qy 528 CGCAACATCACAAACAGCAGCCAGTCCCAAGATAGACGTGTCATAGCAAGGAGCAACGAGT 587
 Db 960 CGCAACATCACAAACAGCAGCCAGTCCCAAGATAGACGTGTCATAGCAAGGAGCAACGAGT 1019
 Qy 588 GCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTCTCTCCCTCGCTGTGAAG 647
 Db 1020 GCAGCTGAAAAGCCATCAGTGTGCACTCCACCCCTGAGGGGTCTCTCCCTCGCTGTGAAG 1079
 Qy 648 ATGATCTTGGAGATTATGCATAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTGCC 707
 Db 1080 ATGATCTTGGAGATTATGCATAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTGCC 1139
 Qy 708 CTGAAGATCTGGGCCATAAATCTTGTAGGGCTCTCATGGCAAGGAGGAGCGAAC 767
 Db 1140 CTGAAGATCTGGGCCATAAATCTTGTAGGGCTCTCATGGCAAGGAGGAGCGAAC 1199
 Qy 768 CTGAAGAGGTAGAGCAAGATACCGAGCAAAATCACCATCTCTCTGCTGCAAGACCTT 827
 Db 1200 CTGAAGAGGTAGAGCAAGATACCGAGCAAAATCACCATCTCTCTGCTGCAAGACCTT 1259
 Qy 828 ACCCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGATTTGTCAGG 887
 Db 1260 ACCCTTTACACCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGATTTGTCAGG 1319
 Qy 888 GCGAGCAGGAAATAATGAAGAAAGTTCGGGAGGCTATGAGAAATGATGCTGCCATG 947
 Db 1320 GCGAGCAGGAAATAATGAAGAAAGTTCGGGAGGCTATGAGAAATGATGCTGCCATG 1379
 Qy 948 AGC-----TCTCACCTGATCCCTGGCCTGAACTGGCTGTGTAGGTCTTTTCCAGCT 1001
 Db 1380 AGCCTGCACTCTCACCTGACCCCTGGCCTGAACTGGCTGTGTAGGTCTTTTCCAGCT 1439
 Qy 1002 TCATCCAGGAGTCCCGCGCTCCAGCAGCTTACTGGGCTGCTCCCTATAGTCTC 1061
 Db 1440 TCATCCAGGAGTCCCGCGCTCCAGCAGCTTACTGGGCTGCTCCCTATAGTCTC 1499
 Qy 1062 TTTATGAGGCTCCCGAGCAGAGATGGTGAGGTGTTTATCCCGCCCGAGGAGTGGG 1121
 Db 1500 TTTATGAGGCTCCCGAGCAGAGATGGTGAGGTGTTTATCCCGCCCGAGGAGTGGG 1559
 Qy 1122 GCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTCCAGCGCTCC 1181
 Db 1560 GCCATCATCGGCAAGAGGGGAGCAGCATCAACAGCTCTCCCGGTTTCCAGCGCTCC 1619
 Qy 1182 ATCAGATTCACACCGAGGACAGCTGACTCCAAAGTTCGATGTTATCATCTCATGGA 1241
 Db 1620 ATCAGATTCACACCGAGGACAGCTGACTCCAAAGTTCGATGTTATCATCTCATGGA 1679
 Qy 1242 CCGCAGAGGCCCCAATTCAGAGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1301
 Db 1680 CCGCAGAGGCCCCAATTCAGAGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAAC 1739

Qy 1302 TTCTTTGTTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTTGTCAGCATCAGCA 1361
 Db 1740 TTCTTTGTTCCCAAGGAGGAGTGAAGCTGGAGACCCACATACGTTGTCAGCATCAGCA 1799
 Qy 1362 GCTGCGCGGGTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTTCAGAAATTTGACGCA 1421
 Db 1800 GCTGCGCGGGTCATTGGCAAAAGTGGAAAAACGGTGAACGAGTTTCAGAAATTTGACGCA 1859
 Qy 1422 GCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACAGCAGAGTCACTGTA 1481
 Db 1860 GCTGAGGTGTAGTACCAAGAGACAGACCCCTGATGAGAACAGCAGAGTCACTGTA 1919
 Qy 1482 ATCATCGACATTTCTATGCCAGTCAGATGCTCAACGAGATCCGAGACATCTGGCC 1541
 Db 1920 ATCATCGACATTTCTATGCCAGTCAGATGCTCAACGAGATCCGAGACATCTGGCC 1979
 Qy 1542 CAGGTTAAGCAGCAGCATCAGAAAGGACAGAGTAACAGGGCCAGCGACCGAGGAAGTGA 1601
 Db 1980 CAGGTTAAGCAGCAGCATCAGAAAGGACAGAGTAACAGGGCCAGCGACCGAGGAAGTGA 2039
 Qy 1602 CCAGCCCTCCCTGTCCTTNGAGTCCAGGACCAACACGGGAGAAATCGAGAGTGTGCT 1661
 Db 2040 CCAGCCCTCCCTGTCCTTNGAGTCCAGGACCAACACGGGAGAAATCGAGAGTGTGCT 2099
 Qy 1662 CTCGCCGAGCCCTGAGAAATGAGTGGGAATCCGGGACACNTGGCCGCGCTGTAGATCA 1721
 Db 2100 CTCGCCGAGCCCTGAGAAATGAGTGGGAATCCGGGACACCTGGGCGCGCTGTAGATCA 2159
 Qy 1722 GGTTCGCCACCTTGATTCAGAAAGATGTTCCAGTCAGGAACCTCTGATCTNTCAG-CCCCA 1780
 Db 2160 GGTTCGCCACCTTGATTCAGAAAGATGTTCCAGTCAGGAACCTCTGATCTNTCAG-CCCCA 2219
 Qy 1781 AACACCCACCAATTTGGCCCAACACTGNTGCCCCCTCGGGGTGTGAGAAATTTAGCGCA 1840
 Db 2220 AACACCTACCAATTTGGCCCAACACTGNTGCCCCCTCGGGGTGTGAGAAATTTAGCGCA 2279
 Qy 1841 AGGCATTTTAAACGTCGATTTTAAAGAGAGCTCTCCAGGCCCCACCAAGAGGTTGAT 1900
 Db 2280 AGGCATTTTAAACGTCGATTTTAAAGAGAGCTCTCCAGGCCCCACCAAGAGGTTGAT 2339
 Qy 1901 CACACCTCAGTGGGAGAAATAAATAATTTCTTCAGGTTTTT 1942
 Db 2340 CACACCTCAGTGGGAGAAATAAATAATTTCTTCAGGTTTTT 2381

RESULT 6
 AR171864
 LOCUS AR171864 1708 bp DNA linear PAT 17-DEC-2001
 DEFINITION Sequence 5 from patent US 6297364.
 ACCESSION AR171864
 VERSION AR171864.1 GI:17910814
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 UNCLASSIFIED.
 REFERENCE 1 (bases 1 to 1708)
 AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K. and Old,L.J.
 TITLE Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof
 JOURNAL Patent: US 6297364-A 5 02-OCT-2001;
 FEATURES Location/Qualifiers
 source 1..1708
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 85.8%; Score 1669.2; DB 6; Length 1708;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 268 CCAGGCAAGCCATCATGAAGCTGAAGTGGCCACCAAGTGGAGAACCATGCCCTGAAGGTCT 327

```
QY 768 CTGAAGAAGGTAGAGCAAGATACCGGACAAAGAAATCACTATCTCCGTCGTTGCAAGACCTT 827
Db 904 CTGAAGAAGGTAGAGCAAGATACCGGACAAAGAAATCACTATCTCCGTTGCAAGACCTT 963
QY 828 ACCCTTTACAACCTCTGAGAGGACCATCACTGTCAAGGGGGCCATCGAGAAATGCTGCAGG 887
Db 964 ACCCTTTACAACCTCTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATGCTGCAGG 1023
QY 888 GCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGTGCTGCGCATG 947
Db 1024 GCCGAGCAGGAAATAATGAAGAAAGTTCCGGAGGCGCTATGAGAATGATGTGCTGCGCATG 1083
QY 948 AGC-----TCTCACTGATCCCTGGCTGAACCTGGCTGCTAGTGTCTTTTCCAGCT 1001
Db 1084 AGCCTGCAGTCTCACTGATCCCTGGCTGAACCTGGCTGCTAGTGTCTTTTCCAGCT 1143
QY 1002 TCATCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTCC 1061
Db 1144 TCATCAGGCGAGTCCCGCGCTCCAGCAGCGTTACTGGGCTGCTCCCTATAGCTCC 1203
QY 1062 TTTATCAGGCTCCGAGCAGGAGATGTTGTCAGGTGTTATCCCGCCCGCAGGCGAGTGGC 1121
Db 1204 TTTATCAGGCTCCGAGCAGGAGATGTTGTCAGGTGTTATCCCGCCCGCAGGCGAGTGGC 1263
QY 1122 GCCATCATCGGCAAGAGGGGCGAGCAGATCAAAAGCTCTCCCGGTTTGCAGGCGCTCC 1181
Db 1264 GCCATCATCGGCAAGAGGGGCGAGCAGATCAAAAGCTCTCCCGGTTTGCAGGCGCTCC 1323
QY 1182 ATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGCTATGTTATCATCACTGGA 1241
Db 1324 ATCAAGATTGCACACCCGAAACACCTGACTCCAAAGTTGCTATGTTATCATCACTGGA 1383
QY 1242 CGCCAGAGGCCCAATTCAGGCTCAGGAGAGAAATCTATGGCAAACTCAAGAGGAGAAC 1301
Db 1384 CGCCAGAGGCCCAATTCAGGCTCAGGAGAGAAATCTATGGCAAACTCAAGAGGAGAAC 1443
QY 1302 TTTCTTTGTTCCCAAGAGGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCA 1361
Db 1444 TTTCTTTGTTCCCAAGAGGAGTGAAGTGGAGACCCACATACGTGTGCCAGCATCAGCA 1503
QY 1362 GTGGCGGGTCAATTGCAAGGTGAAACCGTGAAACCGTGAACGAGTTGAGAAATTTGACGGCA 1421
Db 1504 GTGGCGGGTCAATTGCAAGGTGAAACCGTGAAACCGTGAACGAGTTGAGAAATTTGACGGCA 1563
QY 1422 GCTGAGGTGTTAGTACCAAGAGACCAAGACCCCTGATGAGAAACCAAGAGTCACTGTGAAA 1481
Db 1564 GCTGAGGTGTTAGTACCAAGAGACCAAGACCCCTGATGAGAAACCAAGAGTCACTGTGAAA 1623
QY 1482 ATCATCGGACATTTCTATGCTCAGTCCAGTGGCTCAAGGGAAGATCCGAGACATCTGGCC 1541
Db 1624 ATCATCGGACATTTCTATGCTCAGTGGCTCAAGGGAAGATCCGAGACATCTGGCC 1683
QY 1542 CAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAAACAGGCGCCAGGCAAGGAGAAAGTGA 1601
Db 1684 CAGGTTAAGCAGCAGCATCAGAGGAGCAGAGTAAACAGGCGCCAGGCAAGGAGAAAGTGA 1743
QY 1602 CAGGCGCTCCCTGCTCCCTTNGAGTCCAGAGCAACAAACGGGAGAAATCGAGAGTGTGCT 1661
Db 1744 CAGGCGCTCCCTGCTCCCTTNGAGTCCAGAGCAACAAACGGGAGAAATCGAGAGTGTGCT 1803
QY 1662 CTCCTCCGCGAGGCTCAGAGATGATGGGAATCCGGGACACNTGGCGGCTGTAGATCA 1721
Db 1804 CTCCTCCGCGAGGCTCAGAGATGATGGGAATCCGGGACACNTGGCGGCTGTAGATCA 1863
QY 1722 GGTTCGCCCACTTGAATGAGAAAGATGTTCCAGTGAAGAACCTGTATCTTCAGGCCCAA 1781
Db 1864 GGTTCGCCCACTTGAATGAGAAAGATGTTCCAGTGAAGAACCTGTATCTTCAGGCCCAA 1923
QY 1782 ACACCCCAACCAATGGCCCAACCACTGTNTGCCCTCTGGGGTGTGAGAAATTTAGCGCAA 1841
Db 1924 ACACCCCAACCAATGGCCCAACCACTGTNTGCCCTCTGGGGTGTGAGAAATTTAGCGCAA 1983
```

```
QY 1842 GGCACCTTTAAACGTGATGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATC 1901
Db 1984 GGCACCTTTAAACGTGATGTTTAAAGAGCTCTCCAGGCCCCACCAAGAGGTTGATC 2043
QY 1902 ACACCTCAGTGGGAGAGAAAAATAAATTTCTTCCAGTTTAAAA 1946
Db 2044 ACACCTCAGTGGGAGAGAAAAATAAATTTCTTCCAGTTTAAAA 2088

RESULT 5
AF198254
LOCUS 2381 bp mRNA linear PRI 02-MAR-2000
DEFINITION Homo sapiens mRNA-binding protein CRDBP mRNA, complete cds.
ACCESSION AF198254
VERSION AF198254.1 GI:7141071
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2381)
AUTHORS Ioannidis,P., Trangas,T., Dimitriadis,E., Samiotaki,M.,
Panoutsakopoulos,G., Kyriazoglou,I., Voutzoulis,S.,
Tsiapalis,C.M., Kittas,C., Agnantis,N. and Pandis,N.
TITLE Ectopic expression of a KH-domain containing protein, highly
homologous to both human IMP-1 and mouse CRD-BP, in benign and
malignant mesenchymal tumors
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 2381)
TITLE Ioannidis,P., Trangas,T., Dimitriadis,E. and Samiotaki,M.
JOURNAL Direct Submission
JOURNAL Submitted (25-OCT-1999) G. Papanikolaou Research Center, St Savvas
Hospital, 171 Alexandras Avenue, Athens, Attica 11522, Greece
FEATURES
Location/Qualifiers
1..2381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21.3-Q22"
/cell_line="MCF-7"
306..2039
/note="KH-domain containing protein; similar to Mus
musculus CRD-BP (c-myc mRNA coding region
determinant-binding protein) and Homo sapiens IMP-1
(insulin growth factor II mRNA-binding protein)"
/codon_start=1
/product="mRNA-binding protein CRDBP"
/protein_id="AAF37203.1"
/db_xref="GI:7141072"
/translation="MNKLYIGNESVTPADLEKVPFAEHKISYSGFLVKSGYAFVDC
PDEHMAKAEITFSKVELQGRLEIHSVPKQRSKTQIRNIPQLRWELDSLILA
QYGVENCQVNTSETVAVNVVYSNREQTRQAIIMKLNHQLHNHALKVSYTPDEQIA
QGPENRRGGFGSRGQPGQSPVAAGAPAKQQVDIPLRLVPTQVVGAIIGKEGATI
RNIKTQSKIDVHRKENAGAAEKASIVHSTPEGSSACKMLEIMHKEAKTKTDAE
VPLKTLANNFVRLIKEGRNKKVEODTETKITISLQDLTLNPERTITVKGAIIE
NCRABEIMKKVREAYENDVAMSLQSHLTFLGLNLAAGLFPASSAVPPPSVTVG
AAPYSFMQAPQEMVQVPIPAQVAGIIGKQHQIKQUSRFASASIKIAPPETPDSK
VMVITIGPEAQOQRIYGKJKEENFFGPEKVEKLETHIRVPASAGRIVIGKGGK
TVNEQLNTAAEWWVVRDQTPDNDQVIVKIIFHYASQMAQRKIRIDILAQVKQHQK
GQSNQAQARK"

Query Match 92.3%; Score 1796.6; DB 9; Length 2381;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1826; Conservative 0; Mismatches 9; Indels 7; Gaps 2;

QY 108 AGAGCGCGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGAC 167
Db 540 AGAGCGCGAAATTCAAATCCGAATATTCACCCAGCTCCGATGGGAGTACTGGAC 599
QY 168 AGCCTCTCGCTCAGTATGTTAGTAGAAGTGTGAGCAAGTGAACACCGAGAGTGAG 227
```


Db	1801	AAACCTGTNTGCCCTCGGGGTCTCAGAAATTTNAGCGCAAGCAGCTTTTAAACGTGGAT	1860	Db	181	AGTATGGTACAGTAGAGAACTGTGACCAAGTGAACACCGAGAGTGAGACGGCAGTGGTGA	240
QY	1861	TGTTTAAAGAAGCTCTCCAGGCCCCCAAGAGGGTGGATCACCACCTCAGTGGGAAGAAA	1920	QY	241	ATGTCACTTATCCAAACCGGGAGACAGCAGCAAGCCATCATGAAGCTGAATGGCCACC	300
Db	1861	TGTTTAAAGAAGCTCTCCAGGCCCCCAAGAGGGTGGATCACCACCTCAGTGGGAAGAAA	1920	Db	241	ATGTCACTTATCCAAACCGGGAGACAGCAGCAAGCCATCATGAAGCTGAATGGCCACC	300
QY	1921	AATAAAATTCCTTCAGGTTTAAAA	1946	QY	301	AGTTGGAGAACCATGCGCTGAAGGTCTCTCATATCCCGATGAGCAGATAGCAGGGAC	360
Db	1921	AATAAAATTCCTTCAGGTTTAAAA	1946	Db	301	AGTTGGAGAACCATGCGCTGAAGGTCTCTCATATCCCGATGAGCAGATAGCAGGGAC	360
RESULT 3				QY	361	CTGAGAAATGGCGCGCAGAGGGGCTTTGGCTCTCGGGGTGAGCCGCCAGGGCTCACCTG	420
LOCUS	BD209926			Db	361	CTGAGAAATGGCGCGCAGAGGGGCTTTGGCTCTCGGGGTGAGCCGCCAGGGCTCACCTG	420
DEFINITION	Isolated nucleic acid molecules encoding cancer-associated			QY	421	TGSCAGCGGGGGCCCCCAGCAGCAAGTGGACATCCCCCTTCGGCTCTCTGGTGC	480
ACCESSION	BD209926			Db	421	TGSCAGCGGGGGCCCCCAGCAGCAAGTGGACATCCCCCTTCGGCTCTCTGGTGC	480
VERSION	BD209926.1	GI:33019696		QY	481	CCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAA	540
KEYWORDS	JP 2002512049-A/5.			Db	481	CCACCCAGTATGTGGGTGCCATTATTGGCAAGAGGGGGCCACCATCCGCAACATCACAA	540
SOURCE	Homo sapiens (human)			QY	541	AACAGACCCAGTCCCAAGATAGACGTGCATAGAAAGGAGAACGACGTGACAGCTGAAAAAG	600
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			Db	541	AACAGACCCAGTCCCAAGATAGACGTGCATAGAAAGGAGAACGACGTGACAGCTGAAAAAG	600
REFERENCE	1 (bases 1 to 1946)			QY	601	CCATCAGTGTGCATCCACCCCTGAGGGGTGCTCTCTCGCTTGTAAAGATCATCTTGAGA	660
AUTHORS	Chen,Y.T., Gure,A., Tsang,S., Stockert,E., Jager,E., Knuth,A. and Old,J.J.			Db	601	CCATCAGTGTGCATCCACCCCTGAGGGGTGCTCTCTCGCTTGTAAAGATCATCTTGAGA	660
TITLE	Isolated nucleic acid molecules encoding cancer-associated			QY	661	TTATGATTAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTCCCTTGAAGATCCTGG	720
JOURNAL	Patent: JP 2002512049-A 5 23-APR-2002;			Db	661	TTATGATTAAGAGGCTAAGGACACCAAAACGGCTGACGAGTTCCCTTGAAGATCCTGG	720
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH			QY	721	CCATAATAACTTTGTAGGGCGTCTCATTTGGCAAGAGGACCGAAACCTGAAAGAGTAG	780
	OS Homo sapiens (human)			Db	721	CCATAATAACTTTGTAGGGCGTCTCATTTGGCAAGAGGACCGAAACCTGAAAGAGTAG	780
	PN JP 2002512049-A/5			QY	781	AGCAAGATACCGAGACAAAATAACCATCTCTCTGTTGCAAGACCTTACCTTTTACAACC	840
	PR 16-MAR-1999 JP 2000545030			Db	781	AGCAAGATACCGAGACAAAATAACCATCTCTCTGTTGCAAGACCTTACCTTTTACAACC	840
	PI YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE			QY	841	CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTCAGGGCGGACAGAAA	900
	PI JAGER			Db	841	CTGAGAGGACCATCACTGTGAAGGGGGCCATCGAGAAATTTGTCAGGGCGGACAGAAA	900
	PI ALEXANDER KNUTH,LLOYD J OLD			QY	901	TAATGAAGAAAGTTCCGGAGGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGA	960
	PC C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,			Db	901	TAATGAAGAAAGTTCCGGAGGGCTATGAGAATGATGTGGCTGCCATGAGCTCTCACCTGA	960
	PC C12N1/15,			QY	961	TCCCTGGCTGAACTCTGGTGTGTAGTCTTTTCCAGCTTATCCAGGCTTATCCAGGCTCCGC	1020
	PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/			Db	961	TCCCTGGCTGAACTCTGGTGTGTAGTCTTTTCCAGCTTATCCAGGCTTATCCAGGCTCCGC	1020
	574,			QY	1021	CGCTCCAGCAGCGTTACTGGGGTGTCTTCCAGCTTATCCAGGCTTATCCAGGCTCCGCGC	1080
	PC C12N15/00,C12N5/00			Db	1021	CGCTCCAGCAGCGTTACTGGGGTGTCTTCCAGCTTATCCAGGCTTATCCAGGCTCCGCGC	1080
	CC n=not determined(1622,1702,1771,1809,1833)			QY	1081	AGGAGATGTCGAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATCGGCAAGAGG	1140
	PH Key Location/Qualifiers			Db	1081	AGGAGATGTCGAGGTGTTTATCCCGCCAGCAGTGGCGCCATCATCGGCAAGAGG	1140
	FT source 1..1946			QY	1141	GGCAGCACATCAACAGCTCTCCGGTTTGGCAGCGCTCCATCAAGATTGACACCCG	1200
	FT Location/Qualifiers			Db	1141	GGCAGCACATCAACAGCTCTCCGGTTTGGCAGCGCTCCATCAAGATTGACACCCG	1200
	FT /organism='Homo sapiens (human)'			QY	1201	AAACCTCAGTCCCAAGTTCGATGGTATCATCTGAGCGCGCTCCATCAAGATTGACACCCG	1260
	FT 1..1946			Db	1201	AAACCTCAGTCCCAAGTTCGATGGTATCATCTGAGCGCGCTCCATCAAGATTGACACCCG	1260
	FT /organism='Homo sapiens'			QY	1261	AGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAG	1320
	FT /mol_type='genomic DNA'			Db	1261	AGGCTCAGGGAAGAAATCTATGGCAAACTCAAGGAGGAGAACTTCTTTGGTCCCAAGGAG	1320
	FT /db_xref='taxon:9606'			QY			
FEATURES	source			Db			
ORIGIN							
Query Match	99.7%; Score 1941; DB 6; Length 1946;						
Best Local Similarity	100.0%; Pred. No. 0;						
Matches 1946; Conservative	0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	GCTGTAGCGAGGGGTGGGGGGTCTGTCTGTCTCCCTTCTTCTGCGCGCTCGGGCTCAGC	60	QY	1	GCTGTAGCGAGGGGTGGGGGGTCTGTCTGTCTCCCTTCTTCTGCGCGCTCGGGCTCAGC	60
Db	1	GCTGTAGCGAGGGGTGGGGGGTCTGTCTGTCTCCCTTCTTCTGCGCGCTCGGGCTCAGC	60	Db	1	GCTGTAGCGAGGGGTGGGGGGTCTGTCTGTCTCCCTTCTTCTGCGCGCTCGGGCTCAGC	60
QY	61	CCACCCAGAGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGGTTAGGAGCGGAAAA	120	QY	61	CCACCCAGAGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGGTTAGGAGCGGAAAA	120
Db	61	CCACCCAGAGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGGTTAGGAGCGGAAAA	120	Db	61	CCACCCAGAGCGGGGTGGAGGGGAGTGTCTCAGCTTCCCGGGTTAGGAGCGGAAAA	120
QY	121	TTCAATCCGAATATTCACCCAGCTCCGATGGGAGTCTGGACAGCTCTGCTC	180	QY	121	TTCAATCCGAATATTCACCCAGCTCCGATGGGAGTCTGGACAGCTCTGCTC	180
Db	121	TTCAATCCGAATATTCACCCAGCTCCGATGGGAGTCTGGACAGCTCTGCTC	180	Db	121	TTCAATCCGAATATTCACCCAGCTCCGATGGGAGTCTGGACAGCTCTGCTC	180
QY	181	AGTATGTTACAGTAGAACTGTGAGCAAGTGAACCCGAGAGTGAGACGGCAGTGGTGA	240	QY	181	AGTATGTTACAGTAGAACTGTGAGCAAGTGAACCCGAGAGTGAGACGGCAGTGGTGA	240
Db	181	AGTATGTTACAGTAGAACTGTGAGCAAGTGAACCCGAGAGTGAGACGGCAGTGGTGA	240	Db	181	AGTATGTTACAGTAGAACTGTGAGCAAGTGAACCCGAGAGTGAGACGGCAGTGGTGA	240

LOCUS AR343075 1946 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 7 from patent US 6576756.
ACCESSION AR343075
VERSION AR343075.1 GI:33738477
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1946)
AUTHORS Chen, Y.-T., Gure, A., Teang, S., Stockert, E., Jager, E., Alexander, K. and Old, L. J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof
JOURNAL Patent: US 6576756-A 7 10-JUN-2003;
FEATURES
source location/Qualifiers
1..1946
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.7%; Score 1941; DB 6; Length 1946;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCTGTAGCGAGGGGCTGGGGGGCTGCTCTGTCCCTTCCTTGGCGCTGGCGCTCAGC 60
Db 1 GCTGTAGCGAGGGGCTGGGGGGCTGCTCTGTCCCTTCCTTGGCGCTGGCGCTCAGC 60
Qy 61 CCACCCAGAGGCCGGGGTGGGAGGCGAGTCTCAGCTTCCCGGGTTAGGAGCGGAAAA 120
Db 61 CCACCCAGAGGCCGGGGTGGGAGGCGAGTCTCAGCTTCCCGGGTTAGGAGCGGAAAA 120
Qy 121 TTCAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCCTGTGGCTC 180
Db 121 TTCAATCCGAAATATTCCACCCAGCTCCGATGGGAAGTACTGGACAGCCTGTGGCTC 180
Qy 181 AGTATGGTACGTAGAGACTGTGAGGAAGTGAACACCGAGAGTGAAGCGGAGTGTGA 240
Db 181 AGTATGGTACGTAGAGACTGTGAGGAAGTGAACACCGAGAGTGAAGCGGAGTGTGA 240
Qy 241 ATGTCACTATTCCAAACCGGAGCAGACACGAGGAGCCATCATGAAGCTGAATGCCACC 300
Db 241 ATGTCACTATTCCAAACCGGAGCAGACACGAGGAGCCATCATGAAGCTGAATGCCACC 300
Qy 301 AGTTGGAGAACCATGCCCTGAAGGTCTCTTACATCCCGGATGAGCAGATAGCACAGGAC 360
Db 301 AGTTGGAGAACCATGCCCTGAAGGTCTCTTACATCCCGGATGAGCAGATAGCACAGGAC 360
Qy 361 CTGAGATGGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGGCTCACCTG 420
Db 361 CTGAGATGGGGCGCGAGGGGCTTTGGCTCTCGGGGTGAGCCCGCAGGGCTCACCTG 420
Qy 421 TGGCAGCGGGGCCCCAGCAGCAGCAGTGGACATCCCGCTTCGGCTCTCTGGTGC 480
Db 421 TGGCAGCGGGGCCCCAGCAGCAGCAGTGGACATCCCGCTTCGGCTCTCTGGTGC 480
Qy 481 CCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCAAA 540
Db 481 CCACCCAGTATGTGGTGCCATTATTGGCAAGGAGGGGCCACCATCCGCAACATCAAA 540
Qy 541 AACAGACCCAGTCCAGATAGACGTGCATAGGAAGAGAACCGCAGTGCAGTGAAGAAG 600
Db 541 AACAGACCCAGTCCAGATAGACGTGCATAGGAAGAGAACCGCAGTGCAGTGAAGAAG 600
Qy 601 CCATCAGTGTGCATCCACCCCTGAGGGTGTCTCTCGCTTGTAAAGTATCTTGGAGA 660
Db 601 CCATCAGTGTGCATCCACCCCTGAGGGTGTCTCTCGCTTGTAAAGTATCTTGGAGA 660
Qy 661 TTATGCATAAAGAGCTTAAGGACACCAAAAAGGCTGACGAGTTCCCTTGAAGATCTTGG 720
Db 661 TTATGCATAAAGAGCTTAAGGACACCAAAAAGGCTGACGAGTTCCCTTGAAGATCTTGG 720
Qy 721 CCCATAATAACTTTGTAGGGCGTCTCATTTGGCAAGGAGGAGCGAACCTGAAGAAGTAG 780

Db 721 CCCATAATAACTTTGTAGGGCGTCTCATTTGCAAGGAGGAGCGAACCTGAAGAAGTAG 780
Qy 781 AGCAAGATACCGAGACAAAAATCCATCTCTCTGTTGCAAGACCTTACCTTTTCAACC 840
Db 781 AGCAAGATACCGAGACAAAAATCCATCTCTCTGTTGCAAGACCTTACCTTTTCAACC 840
Qy 841 CTGAGAGACCATCACTGTGAAGGGGCCATCGAGAATTGTCAGGGCCGAGCAGGAAA 900
Db 841 CTGAGAGACCATCACTGTGAAGGGGCCATCGAGAATTGTCAGGGCCGAGCAGGAAA 900
Qy 901 TAATGAAGAAAGTTTCGGAGGCCCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCTGA 960
Db 901 TAATGAAGAAAGTTTCGGAGGCCCTATGAGAAATGATGTGGCTGCCATGAGCTCTCACCTGA 960
Qy 961 TCCCTGGCTGAACCTGGCTCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGC 1020
Db 961 TCCCTGGCTGAACCTGGCTCTGTAGGTCTTTTCCAGCTTCATCCAGCGCAGTCCCGC 1020
Qy 1021 CGCTCCAGCAGCGTTACTTGGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCCGAGC 1080
Db 1021 CGCTCCAGCAGCGTTACTTGGGGCTGCTCCCTATAGCTCCTTTATGAGGCTCCCGAGC 1080
Qy 1081 AGGAGATGGTGCAGGTGTTTATCCCGCCAGGCGAGTGGGGCCCATCATCGGCAAGAAG 1140
Db 1081 AGGAGATGGTGCAGGTGTTTATCCCGCCAGGCGAGTGGGGCCCATCATCGGCAAGAAG 1140
Qy 1141 GGCAGCACATCAAAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCCCG 1200
Db 1141 GGCAGCACATCAAAACAGCTCTCCCGGTTTGCAGCGCTCCATCAAGATTGCACCCCG 1200
Qy 1201 AAACACCTGACTCCAAAGTTCTGTTATCTACTGGACCGCGAGGGCCCAATTCA 1260
Db 1201 AAACACCTGACTCCAAAGTTCTGTTATCTACTGGACCGCGAGGGCCCAATTCA 1260
Qy 1261 AGGCTCAGGGAGAAATCTATGGCAAACTCAAGGAGGAACTTCTTTGGTCCCAAGGAG 1320
Db 1261 AGGCTCAGGGAGAAATCTATGGCAAACTCAAGGAGGAACTTCTTTGGTCCCAAGGAG 1320
Qy 1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCGGGTCAATTGGA 1380
Db 1321 AAGTGAAGCTGGAGACCCACATACGTGTGCCAGCATCAGCAGCTGGCGGGTCAATTGGA 1380
Qy 1381 AAGTGGAAAAACGGTGAACCGAGTTGAGAAATTTACCGCAGCTGAGGTGGTAGTACCAA 1440
Db 1381 AAGTGGAAAAACGGTGAACCGAGTTGAGAAATTTACCGCAGCTGAGGTGGTAGTACCAA 1440
Qy 1441 GAGACCAACCCCTGATCAGAAACGACAGGTCATCGTGAATAATCATCGACATTTCTATG 1500
Db 1441 GAGACCAACCCCTGATCAGAAACGACAGGTCATCGTGAATAATCATCGACATTTCTATG 1500
Qy 1501 CCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCGAGGTTAAGCAGCAGCATC 1560
Db 1501 CCAGTCAGATGGCTCAACGGAAGATCCGAGACATCCTGGCCGAGGTTAAGCAGCAGCATC 1560
Qy 1561 AGAAGGACAGAGTAACACCGCCAGCAGCGAGGAAGTGACCGCCCTCTCTCCCT 1620
Db 1561 AGAAGGACAGAGTAACACCGCCAGCAGCGAGGAAGTGACCGCCCTCTCTCCCT 1620
Qy 1621 TNGAGTCCAGACAAACACGGCGAGAAATCGAGAGTGTCTCTCCCGCAGCGCTGAGA 1680
Db 1621 TNGAGTCCAGACAAACACGGCGAGAAATCGAGAGTGTCTCTCCCGCAGCGCTGAGA 1680
Qy 1681 ATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGGCCATTGATGA 1740
Db 1681 ATGAGTGGGAATCCGGGACACNTGGGCGGGCTGTAGATCAGGTTTGGCCATTGATGA 1740
Qy 1741 GAAAGATCTTCAGTGAAGAACCCCTGATCTNTCAGCCCCAACACCCCAATTTGGCC 1800
Db 1741 GAAAGATCTTCAGTGAAGAACCCCTGATCTNTCAGCCCCAACACCCCAATTTGGCC 1800
Qy 1801 AACACTGNTGCCCCCTCGGGGTGTGAGAAATTTAGCGCAAGGCACTTTTAAACGTTGAT 1860

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 03:38:17 ; Search time 7493.35 Seconds
(without alignments)
11256.058 Million cell updates/sec

Title: US-09-270-437D-7
Perfect score: 1946
Sequence: 1 gctgtacggagggtctggg.....attcccttcagggttttaaaa 1946

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vl.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1941	99.7	1946	6	AR171866	AR171866 Sequence
2	1941	99.7	1946	6	AR343075	AR343075 Sequence
3	1941	99.7	1946	6	BD209926	BD209926 Isolated
4	1818	93.4	2130	9	AF117106	AF117106 Homo sapi
5	1796.6	92.3	2381	9	AF198254	AF198254 Homo sapi
6	1669.2	85.8	1708	6	AR171864	AR171864 Sequence
7	1669.2	85.8	1708	6	AR343073	AR343073 Sequence
8	1669.2	85.8	1708	6	BD209924	BD209924 Isolated
9	1389	71.4	2780	6	BD127811	BD127811 Primer fo
10	1389	71.4	2780	9	AK074915	AK074915 Homo sapi
11	1301	66.9	2224	6	AR160244	AR160244 Sequence
12	1298.6	66.7	2223	10	AF061569	AF061569 Mus muscu
13	1297.6	66.7	2444	10	BC051679	BC051679 Mus muscu
14	1250	64.2	2105	10	AF541940	AF541940 Rattus no
15	1046	53.8	2021	5	AF026527	AF026527 Gallus ga
16	732.8	37.7	833	6	BD124998	BD124998 Primer fo
17	732.8	37.7	833	6	BD126684	BD126684 Primer fo
18	723.4	37.2	1740	10	AB046173	AB046173 Mus muscu
19	723.4	37.2	2140	10	BC045138	BC045138 Mus muscu
20	723.4	37.2	4030	10	BC049082	BC049082 Mus muscu
21	674.2	34.6	1740	6	BD275902	BD275902 COMPOUNDS
22	674.2	34.6	1740	6	AR220687	AR220687 Sequence
23	674.2	34.6	1740	6	AR281251	AR281251 Sequence
24	674.2	34.6	1740	6	AX365954	AX365954 Sequence
25	674.2	34.6	1743	6	AX366054	AX366054 Sequence
26	674.2	34.6	1743	6	AX366057	AX366057 Sequence
27	674.2	34.6	1764	6	AX366035	AX366035 Sequence
28	674.2	34.6	4171	9	AF117108	AF117108 Homo sapi
29	672.6	34.6	4155	9	HSU76705	U76705 Human putat
30	672.6	34.6	4159	6	AR171863	AR171863 Sequence
31	672.6	34.6	4159	6	AR343072	AR343072 Sequence
32	672.6	34.6	4159	6	BD209923	BD209923 Isolated
33	672.6	34.6	4181	6	BD275767	BD275767 COMPOUNDS
34	672.6	34.6	4181	6	AR220552	AR220552 Sequence
35	672.6	34.6	4181	6	AR255546	AR255546 Sequence
36	672.6	34.6	4181	6	AR281116	AR281116 Sequence
37	672.6	34.6	4181	6	AX333233	AX333233 Sequence
38	672.6	34.6	4181	6	AX365782	AX365782 Sequence
39	672.6	34.6	4181	9	HSU97188	U97188 Homo sapien
40	672.6	34.6	4602	6	AX397963	AX397963 Sequence
41	667.6	34.3	2011	5	AF161270	AF161270 Danio rer
42	667.6	34.3	2612	5	BC045873	BC045873 Danio rer
43	661.4	34.0	2161	5	AF055923	AF055923 Xenopus l
44	661.4	34.0	2405	5	AF064634	AF064634 Xenopus l
45	650.8	33.4	2031	5	BC057700	BC057700 Xenopus l

ALIGNMENTS

RESULT 1
AR171866
LOCUS AR171866 1946 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 7 from patent US 6297364.
ACCESSION AR171866
VERSION AR171866.1 GI:17910816
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1946)
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof